Wave Optics in Gravitational Lensing

Takahiro T. Nakamura and Shuji Deguchi*

Department of Physics, University of Tokyo, Tokyo 113-0033, Japan
**Nobeyama Radio Observatory, Minamisaku, Nagano 384-1305, Japan*

(Received February 10, 1999)

This review on "wave optics in gravitational lensing" includes a derivation of the diffraction integral formula for the lensed wave amplitude using the path integral (§2), reduction of this formula to the geometric optics approximation in the short wavelength limit along with discussion on the condition that the wave effects become important (§3), examples of wave effects for a point-mass lens and the fold caustic (§4), and a numerical method of evaluating the diffraction integral (§5).

§1. Introduction

Gravitational lensing of astronomical objects causes magnification of their observed images. Since geometric optics provides an excellent approximation in most astrophysical situations for computing this magnification, wave optics, which is more fundamental, is rarely discussed in works on gravitational lensing.

In this paper we review the wave aspects of gravitational lens theory.[1)-16)] The main purpose is to state clearly the conditions under which wave effects are non-negligible. They are negligible in usual situations such as optical lensing of quasars by galaxies, but are non-negligible in such extreme cases as femtolensing of gamma-ray bursts[11),12)] and gravitational lensing of gravitational waves.[16)] More importantly, the geometric optics approximation breaks down near the lens mapping singularity (caustic) where this approximation gives an infinite brightness for a point source. This artifact of infinity should be remedied by application of wave optics theory.

Since wave optics is more fundamental than geometric optics, it is possible to construct gravitational lens theory on the basis of the former *and* with no reference to the latter. So we attempt to make this paper as self-contained as we can. We use the $c = 1$ unit.

§2. Derivation of diffraction integral through the path integral approach

In this section we derive a formula (diffraction integral, Eq. (2·8)) to calculate the amplitude of lensed waves using the path integral approach.*)

2.1. *Monochromatic waves from a point source*

We consider the situation depicted in Fig. 1. Here, a point source of radiation emits spherical, monochromatic waves of frequency ω, which propagate through a

*) To our knowledge this approach is new to the gravitational lens theory.

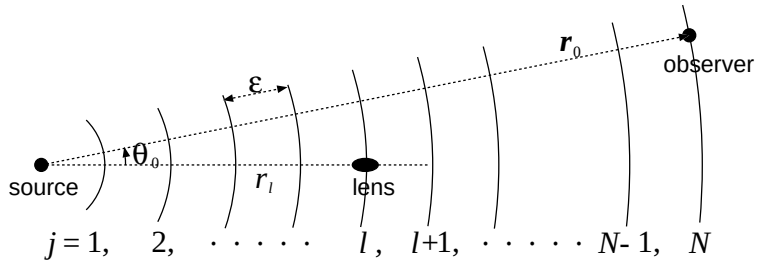

Fig. 1.

gravitational lens and reach a distant observer. The space-time metric

$$ds^2 = -(1+2U)dt^2 + (1-2U)d\vec{r}^{\,2} \tag{2·1}$$

is Minkowskian plus a small disturbance due to the gravitational potential $U(\vec{r})$ ($\ll 1$) of the lensing object. The influence of this potential is localized within a small region around the lens compared with the large distances between source, lens and observer. The propagation equation $\partial_\mu(\sqrt{-g}\, g^{\mu\nu}\partial_\nu\phi) = 0$ of the wave amplitude $\phi(\vec{r},t) = \tilde{\phi}(\vec{r})\, e^{-i\omega t}$ is

$$(\nabla^2 + \omega^2)\tilde{\phi} = 4\omega^2 U\tilde{\phi} \tag{2·2}$$

to first order in U. We use the spherical coordinates (r,θ,φ) with the origin at the source and the polar-axis pointing toward the lens. The observer is located at $\vec{r}_0 = (r_0,\theta_0,\varphi_0)$ with $\theta_0 \ll 1$. Since the waves reaching the observer should be confined in the region of $\theta \ll 1$, we can set $\sin\theta \simeq \theta$ in the equations below and regard $\boldsymbol{\theta} = \theta(\cos\varphi, \sin\varphi)$ as a two-dimensional vector on a flat plane. Without the lensing object, the wave amplitude would be $\tilde{\phi}_0(r) = A\, e^{i\omega r}/r$. We define the amplification factor $F(\vec{r}) = \tilde{\phi}(\vec{r})/\tilde{\phi}_0(r)$ of the wave amplitude due to lensing. Equation (2·2) is rewritten in terms of F as

$$\frac{\partial^2 F}{\partial r^2} + 2i\omega\frac{\partial F}{\partial r} + \frac{1}{r^2}\nabla_\theta^2 F = 4\omega^2 U F\,, \tag{2·3}$$

where $\nabla_\theta^2 = \partial^2/\partial\theta^2 + \theta^{-1}\partial/\partial\theta + \theta^{-2}\partial^2/\partial\varphi^2$. Assuming that $\omega/|\partial \ln F/\partial r| \sim$ (scale on which F varies)/(wavelength) $\gg 1$, we neglect the first term $\partial^2 F/\partial r^2$ compared with the second term (i.e., the eikonal approximation). Then Eq. (2·3) looks like the Schrödinger equation with the "time" coordinate r, the "particle mass" ω, and the "time dependent potential" $2\omega\, U(r,\boldsymbol{\theta})$. The corresponding Lagrangian which yields the classical equation of motion is

$$L(r,\boldsymbol{\theta},\dot{\boldsymbol{\theta}}) = \omega\left[\tfrac{1}{2}r^2|\dot{\boldsymbol{\theta}}|^2 - 2U(r,\boldsymbol{\theta})\right], \tag{2·4}$$

where $\dot{\boldsymbol{\theta}} = d\boldsymbol{\theta}/dr$. From the path integral formulation of quantum mechanics,[17] the solution of Eq. (2·3) is formally written as

$$F(\vec{r}_0) = \int \mathcal{D}\boldsymbol{\theta}(r)\, \exp\left\{ i\int_0^{r_0} dr\, L[r,\boldsymbol{\theta}(r),\dot{\boldsymbol{\theta}}(r)]\right\}. \tag{2·5}$$

This expression contains the following meanings: i) Choose a particular function $\boldsymbol{\theta}(r)$ representing a path from the source at the origin to the observer at $\vec{r}_0 = (r_0, \boldsymbol{\theta}_0)$; ii) Evaluate the integral in the curly brackets along this particular path. The result is the "phase" as a functional of $\boldsymbol{\theta}(r)$; iii) Sum up these phases for all possible functions $\boldsymbol{\theta}(r)$. We divide the distance r_0 between source and observer into infinitesimal "time" steps $\epsilon = r_0/N$ by an infinitely large integer N and specifies the path $\boldsymbol{\theta}(r)$ by the set of coordinates $\boldsymbol{\theta}_j = \boldsymbol{\theta}(r_j)$ $(j = 1, \cdots, N)$ on the j-th sphere of radius $r_j = j\epsilon$ (cf. Fig. 1). Let the lens be located on the l-th sphere (lens plane), so that $r_l = l\epsilon$ is the distance between the source and lens. Here we introduce the "thin lens" approximation, whose formal content is to substitute $\frac{1}{2}\delta(r - r_l)\,\hat{\psi}(\boldsymbol{\theta})$ for $U(r, \boldsymbol{\theta})$ in Eq. (2·5), where

$$\hat{\psi}(\boldsymbol{\theta}) = 2\int_0^{r_0} dr\, U(r, \boldsymbol{\theta}) \tag{2·6}$$

is calculated for fixed $\boldsymbol{\theta}$. Thus Eq. (2·5) is rewritten as

$$F(\vec{r}_0) = \left[\prod_{j=1}^{N-1} \int \frac{d^2\boldsymbol{\theta}_j}{A_j}\right] \exp\left\{ i\omega \left[\epsilon \sum_{j=1}^{N-1} \frac{r_j r_{j+1}}{2} \left| \frac{\boldsymbol{\theta}_j - \boldsymbol{\theta}_{j+1}}{\epsilon} \right|^2 - \hat{\psi}(\boldsymbol{\theta}_l) \right] \right\}, \tag{2·7}$$

where the normalizations $A_j = 2\pi i\epsilon/(\omega\, r_j r_{j+1})$ are such that $F = 1$ if $\hat{\psi} = 0$. The thin lens approximation does not imply that the lens is infinitesimally thin, but that those paths which contribute to the phase integral (Eq. (2·5)) are well approximated by a constant vector $\boldsymbol{\theta}(r) \simeq \boldsymbol{\theta}_l$ within the tiny region of $U(\vec{r}) \neq 0$ compared with the huge distances r_0 and r_l. It is shown in the Appendix that Eq. (2·7) reduces straightforwardly to the desired formula

$$F(\vec{r}_0) = \frac{\omega}{2\pi i} \frac{r_l r_0}{r_{l0}} \int d^2\boldsymbol{\theta}_l \, \exp\left\{ i\omega \left[\frac{r_l r_0}{2 r_{l0}} |\boldsymbol{\theta}_l - \boldsymbol{\theta}_0|^2 - \hat{\psi}(\boldsymbol{\theta}_l) \right] \right\}, \tag{2·8}$$

where r_{l0} is the distance between the lens and observer. Thanks to the thin lens approximation (Eq. (2·6)) we could restrict the possible forms of functions $\boldsymbol{\theta}(r)$ within those paths which go straight from the source to a "deflection point" $\boldsymbol{\theta}_l$ and again go straight from $\boldsymbol{\theta}_l$ to the observer. The first term in the square bracket is the path length difference between a straight path from the source to the observer and a deflected path penetrating the lens plane through $\boldsymbol{\theta}_l$. The second term is the time delay caused by the gravitational field around the lens.

Let us interchange the roles of the source and the observer; namely we set another coordinate system (cf. Fig. 2) in which the origin is at the *observer*, the

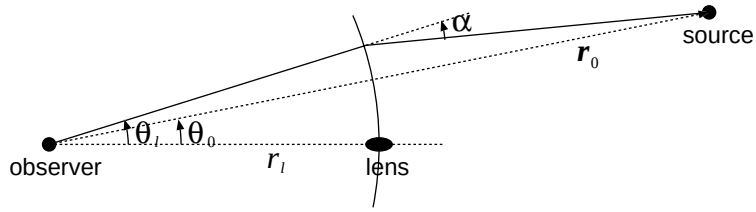

Fig. 2.

polar-axis is pointing toward the lens, and the source is located at $\vec{r}_0 = (r_0, \boldsymbol{\theta}_0)$. If one repeats the above procedure in this coordinate system, then the time delays in the two coordinate systems are identical up to order θ_0^2. Therefore Eq. (2·8) also expresses the lensed wave amplitude for the *source* position $\vec{r}_0$ if r_l is the distance between the observer and lens, and r_{l0} is that between the lens and source.

We rewrite Eq. (2·8) using more compact notation. Introducing the characteristic angle scale θ_* (e.g. Eq. (4·1)) and defining dimensionless quantities

$$\boldsymbol{x} = \boldsymbol{\theta}_l/\theta_*\,, \quad \boldsymbol{y} = \boldsymbol{\theta}_0/\theta_*\,, \quad w = \frac{r_l r_0}{r_{l0}}\,\theta_*^2\,\omega\,, \tag{2·9}$$

$$\psi(\boldsymbol{x}) = \frac{r_{l0}}{r_l r_0}\,\theta_*^{-2}\hat{\psi}(\theta_*\boldsymbol{x})\,, \quad T(\boldsymbol{x},\boldsymbol{y}) = \frac{1}{2}|\boldsymbol{x}-\boldsymbol{y}|^2 - \psi(\boldsymbol{x})\,, \tag{2·10}$$

one obtains

$$F(w,\boldsymbol{y}) = \frac{w}{2\pi i}\int d^2\boldsymbol{x}\,\exp[iw\,T(\boldsymbol{x},\boldsymbol{y})]\,. \tag{2·11}$$

This is the diffraction integral formula expressing the amplification factor of lensed wave amplitudes as a function of the frequency w and the source (or observer) position $\boldsymbol{y}$. The amplification of the wave intensity is $|F|^2$. Hereafter we use the scaled length and time as defined in Eq. (2·9); for example T is simply referred to as the "time delay" instead of the "scaled time delay".

Although the metric in Eq. (2·1) does not take account of the expansion and geometry of our universe, all the formulae in this paper are applicable to cosmological situations, if we use the angular diameter distance for r_0, r_l and r_{l0}, replace ω with $\omega(1+z)$ where z is the lens redshift, and assume that the wavelength is much smaller than the horizon scale. Also the scalar wave analysis in this paper is valid for waves of any helicity, if the rotation of the polarization direction (such as the Faraday rotation of electromagnetic waves) is negligible.

When the lensing object is spherically symmetric, $\psi(\boldsymbol{x})$ depends only on $x = |\boldsymbol{x}|$ and Eq. (2·11) becomes

$$F(w,y) = -iw\,e^{iwy^2/2}\int_0^\infty dx\,x\,J_0(wxy)\,\exp\left\{iw\left[\tfrac{1}{2}x^2 - \psi(x)\right]\right\}\,, \tag{2·12}$$

where J_0 is the Bessel function of zeroth order.

2.2. *Non-monochromatic waves from an extended source*

Let $\tilde{\phi}_0(w,\boldsymbol{y})$ be the amplitude of unlensed waves which are emitted at a point $\boldsymbol{y}$ on the source plane with frequency w. The observed wave amplitude at time τ is the superposition of the lensed one $F\,\tilde{\phi}_0$ over $\boldsymbol{y}$ and w:

$$\phi(\tau) = \int d^2\boldsymbol{y}\int dw\,F(w,\boldsymbol{y})\,\tilde{\phi}_0(w,\boldsymbol{y})\,e^{-iw\tau}\,. \tag{2·13}$$

When the source is observed through a band pass filter, $\tilde{\phi}_0$ in this equation represents the source spectrum at $\boldsymbol{y}$ times the filtering function of frequency. Since $\phi(\tau)$ is real, $\tilde{\phi}_0^*(w) = \tilde{\phi}_0(-w)$. Actually $\tilde{\phi}_0$ should also depend on time because of the fluctuation in the source activity. This fluctuation of $\tilde{\phi}_0$ would be a stochastic process, assuming

that the source activity is stationary in time. So we define the correlation function of the observed wave amplitude as

$$C(\tau) = \frac{\langle \phi(\tau')\, \phi(\tau'+\tau) \rangle}{\langle |\phi(\tau')|^2 \rangle_{F=1}}\,, \tag{2·14}$$

where $\langle \cdots \rangle$ denotes averaging over the ensemble of stochastic functions $\tilde{\phi}_0$. Because of the stationarity of the source activity, $\langle \cdots \rangle$ is also interpreted as an integral with respect to τ' over sufficiently long time $\gg 1/w$. Assuming that the different parts of the source radiate incoherently, one has $\langle \tilde{\phi}_0(w, \boldsymbol{y})\, \tilde{\phi}_0^*(w', \boldsymbol{y}') \rangle = \hat{I}(w, \boldsymbol{y})\, \delta(w - w')\, \delta^2(\boldsymbol{y} - \boldsymbol{y}')$, where $\hat{I}(w, \boldsymbol{y})$ is the source surface brightness (times the filtering function). Therefore one obtains

$$C(\tau) = \int d^2\boldsymbol{y} \int dw\, |F(w, \boldsymbol{y})|^2\, I(w, \boldsymbol{y})\, \cos(w\tau)\,, \tag{2·15}$$

where $I(w, \boldsymbol{y}) = \hat{I}(w, \boldsymbol{y}) / [\int d^2\boldsymbol{y} \int dw\, \hat{I}(w, \boldsymbol{y})]$. The amplification factor of wave intensity is given by $C(0)$.

§3. Geometric optics limit

In this section the formulae derived in the previous section are reduced to the geometric optics approximation in the short wavelength limit, in order to discuss the conditions under which the wave effects are non-negligible.

3.1. *Monochromatic waves from a point source*

Let us imagine a two-dimensional surface $T(\boldsymbol{x}) = \frac{1}{2}|\boldsymbol{x} - \boldsymbol{y}|^2 - \psi(\boldsymbol{x})$ (the time delay function, Eq. (2·10)) above the $\boldsymbol{x}$-plane for a fixed $\boldsymbol{y}$. For example we have $\psi(\boldsymbol{x}) = \ln|\boldsymbol{x}|$ for a point mass lens discussed in §4.1. In the short wavelength limit $w \to \infty$, the integrant of Eq. (2·11) is a rapidly oscillating function and its contribution to the value of F comes from regions near the stationary points of this surface. The stationary points are determined by $\boldsymbol{\partial} T(\boldsymbol{x}) = 0$ or

$$\boldsymbol{y} = \boldsymbol{x} - \boldsymbol{\partial}\psi(\boldsymbol{x}) \quad \text{or} \quad \boldsymbol{\theta}_0 = \boldsymbol{\theta}_l - \frac{r_{l0}}{r_0}\, \boldsymbol{\alpha}(\boldsymbol{\theta}_l)\,, \tag{3·1}$$

where $\boldsymbol{\partial} = \partial/\partial\boldsymbol{x}$ and $\boldsymbol{\alpha}(\boldsymbol{\theta}_l) = r_l^{-1} \partial\hat{\psi}/\partial\boldsymbol{\theta}_l$. These stationary points correspond to the "images" in geometric optics, and the condition $\boldsymbol{\partial} T(\boldsymbol{x}) = 0$ expresses Fermat's principle of least time. In fact Eq. (3·1) is the "lens equation" [10] which determines the positions of images $\boldsymbol{x}$ (or $\boldsymbol{\theta}_l$) for a given source position $\boldsymbol{y}$ (or $\boldsymbol{\theta}_0$) with the deflection angle $\boldsymbol{\alpha}$ of null geodesics (cf. Fig. 2). For a sufficiently "strong" lens potential $\psi(\boldsymbol{x})$, the number of images is more than one. We expand $T(\boldsymbol{x})$ around the j-th image $\boldsymbol{x}_j(\boldsymbol{y})$ as

$$T(\boldsymbol{x}) = T(\boldsymbol{x}_j) + \frac{1}{2} \sum_{ab} \bar{x}_a \bar{x}_b\, \partial_a \partial_b T(\boldsymbol{x}_j) + \frac{1}{6} \sum_{abc} \bar{x}_a \bar{x}_b \bar{x}_c\, \partial_a \partial_b \partial_c T(\boldsymbol{x}_j) + \cdots, \tag{3·2}$$

where $\bar{\boldsymbol{x}} = \boldsymbol{x} - \boldsymbol{x}_j$ and the indices $abc\cdots$ run from 1 to 2. If the frequency w is so large as to satisfy

$$w|\boldsymbol{\partial}^2 T|^3 \gg |\boldsymbol{\partial}^3 T|^2\,, \quad w|\boldsymbol{\partial}^2 T|^2 \gg |\boldsymbol{\partial}^4 T|\,, \quad \cdots \tag{3·3}$$

then the third and higher order terms of this expansion can be neglected in calculating Eq. (2·11) near $\boldsymbol{x}_j$. Assuming that all the images satisfy Eq. (3·3), one obtains using the Gaussian integral*)

$$F = \sum_j |\mu(\boldsymbol{x}_j)|^{1/2} \exp[iw\, T(\boldsymbol{x}_j) - i\pi n_j]\,, \tag{3·4}$$

where

$$\mu(\boldsymbol{x}) = 1/\det[\partial_a\partial_b T(\boldsymbol{x})] = 1/\det[\delta_{ab} - \partial_a\partial_b\psi(\boldsymbol{x})]\,, \tag{3·5}$$

and $n_j = 0,\ 1/2,\ 1$ when $\boldsymbol{x}_j$ is a minimum, saddle, maximum point of $T(\boldsymbol{x})$, respectively. Equation (3·4) implies that the observed wave is a superposition of waves from each image, with the amplitude ratio $|\mu(\boldsymbol{x}_j)|^{1/2}$ and the phase shift $wT(\boldsymbol{x}_j) - \pi n_j$. The amplification of wave intensity is

$$|F|^2 = \sum_j |\mu_j| + 2\sum_{j<k} |\mu_j\,\mu_k|^{1/2} \cos(w\, T_{jk} - \pi n_{jk})\,, \tag{3·6}$$

where $\mu_j = \mu(\boldsymbol{x}_j)$, $T_{jk} = T(\boldsymbol{x}_k) - T(\boldsymbol{x}_j)$ is the arrival time difference between the k-th and j-th images, and $n_{jk} = n_k - n_j$. Note that in geometric optics, the magnification factor of the image size for an infinitesimal source is given as the Jacobian of the lens mapping $\boldsymbol{y} \mapsto \boldsymbol{x}$ in Eq. (3·1):[10]

$$\mu(\boldsymbol{x}) = 1/\det[\boldsymbol{\partial} \otimes \boldsymbol{y}] = 1/\det[\delta_{ab} - \partial_a\partial_b\psi(\boldsymbol{x})]\,. \tag{3·7}$$

From the conservation of surface brightness in geometric optics,[18] this image magnification is equal to the amplification of source intensity. Thus the first term of Eq. (3·6) coincides with the "geometric optics" result, where $|\mu_j|$ is the magnification factor for the j-th image. The second term expresses the "interference" between the images which produces a fringe pattern of wave intensity on the source (or observer) plane (e.g. Fig. 5). Equation (3·4) is the semi-classical approximation in quantum mechanics.

3.2. *Diffraction near lens mapping singularity*

According to geometric optics, the image magnification $|\mu_j|$ for a point source diverges when the image is located near a critical curve, which is a curve on the $\boldsymbol{x}$-plane defined by $\det[\partial_a\partial_b T(\boldsymbol{x})] = 0$ (cf. Eq. (3·5)). A caustic is the corresponding curve on the $\boldsymbol{y}$-plane mapped by Eq. (3·1). (Figure 6 shows an example of a critical curve and a caustic.) Namely, when a point source is located near a caustic, some of the images are near a critical curve and have infinitely large magnifications. Of course this divergence is an artifact of the approximation in Eq. (3·4), because $|\boldsymbol{\partial}^2 T|$ for these images is too small to satisfy Eq. (3·3). Thus the wave amplitude is written as a sum of two terms: $F = F_{\rm c} + F_{\rm d}$. The second term $F_{\rm d}$, the diffraction part, is given by Eq. (2·11) evaluated exactly around these images near the critical curve. This term describes the diffraction phenomenon in the sense that the non-geodesic paths contribute significantly to the wave amplitude. The first term $F_{\rm c}$, the semi-classical part given by Eq. (3·4), is the contribution from other normal images.

*) using $\int_{-\infty}^{\infty} dx\, e^{ix^2} = \sqrt{\pi}\, e^{i\pi/4}$ after diagonalizing $\partial_a\partial_b T$

Equation (3·3) roughly states that the diffraction effect is unimportant when the time delay T_{jk} between images is much larger than the wave period $1/w$. This fact is also seen by the following simple consideration. Regarding the double image in gravitational lensing as the double-slit of Young's interference experiment, the angle of diffraction is $\theta \sim \lambda/d$, where λ is the wavelength and d is distance between the two images. On the other hand, the deflection angle of gravitational lensing is $\alpha \sim d/r_l$ where r_l is the distance from observer to lens. Then the condition $\theta \ll \alpha$ is equivalent to stipulating that the time delay between images $\sim r_l\alpha^2/c$ is much larger than λ/c.

For a spherically symmetric lens, the point $y = 0$ just behind the lens is a caustic producing the "Einstein ring" (e.g. §4.1). The diffraction effect near this caustic is evaluated with Eq. (2·12) for small y. When $w \to \infty$ and $y \ll w^{-1/2}$, Eq. (2·12) yields

$$|F|^2 \simeq 2\pi w\, x_{\rm E}^2\, |1 - \psi''(x_{\rm E})|^{-1}\, J_0^2(w x_{\rm E} y)\,, \tag{3·8}$$

where $x_{\rm E}$ is a (positive) solution of $x = \psi'(x)$. Therefore the maximum value of the amplification for a spherically symmetric lens is given by $2\pi w\, x_{\rm E}^2/|1 - \psi''(x_{\rm E})|$ in the short wavelength limit.

3.3. *Non-monochromatic waves from an extended source*

Let Δy and Δw be the typical sizes of the region where the source surface brightness I is non-zero, and let $\boldsymbol{y}$ and w be the centers of this region; for example, Δy is the radius of a circular source whose center is at $\boldsymbol{y}$, and Δw is the width of a band pass filter whose central frequency is w. For definiteness we assume that the source is sufficiently far from caustics, so that the semi-classical approximation (Eq. (3·6)) is valid in this entire region. Then substituting Eq. (3·6) into Eq. (2·15) results in

$$C(\tau) = {\rm Re}\{e^{-iw\tau}[C_0(\tau) + C_1(\tau) + C_2(\tau)]\}\,, \tag{3·9}$$

where

$$C_0(\tau) = \sum_j \int d^2\boldsymbol{y}'\, |\mu_j(\boldsymbol{y}')|\, \tilde{I}(\tau, \boldsymbol{y}')\,, \tag{3·10}$$

$$C_{1,2}(\tau) = \sum_{j<k} \int d^2\boldsymbol{y}'\, |\mu_j(\boldsymbol{y}')\, \mu_k(\boldsymbol{y}')|^{1/2}\, e^{\pm i[w T_{jk}(\boldsymbol{y}') - \pi n_{jk}]}\, \tilde{I}[\tau \mp T_{jk}(\boldsymbol{y}'),\, \boldsymbol{y}']\,, \tag{3·11}$$

$$\tilde{I}(\tau, \boldsymbol{y}') = \int dw'\, I(w', \boldsymbol{y}')\, e^{-i(w'-w)\tau}\,. \tag{3·12}$$

The terms C_1 and C_2 arise from the interference between images, but the integral in Eq. (3·11) tends to erase the fringe pattern produced by the interference. If the source size Δy is much larger than the typical fringe width $1/|w\, dT_{jk}/dy|$, or if the coherence time $1/\Delta w$ is much smaller than the typical time delay $|\tau \mp T_{jk}(\boldsymbol{y})|$, then this oscillatory integral makes C_1 and C_2 much smaller than C_0. Conversely, when the source size is comparable to or smaller than the fringe width, then it is always possible to choose such a value of τ as to make the interference term C_1 or C_2 observable (e.g. Fig. 4). In the short wavelength limit $w \to \infty$, the fringe width goes to zero and so the interference terms vanish. The geometric optics approximation

uses $C(\tau) = \mathrm{Re}[e^{-iw\tau}\, C_0(\tau)]$ even when the source is located near or straddling a caustic. In this case Eq. (3·10) remains finite though some of $\mu_j(\boldsymbol{y})$ are divergent there. Diffraction near caustics is also unimportant as is interference, when the source size is much larger than the typical fringe width (e.g. Fig. 3).

In order to observe the interference terms C_1 and C_2, not only the source size but also the observer size (or diameter of the telescope) must be smaller than the fringe width. Otherwise further integration of Eq. (3·11) over $\boldsymbol{y}$ smears out the interference pattern. Using the double-slit argument (cf. §3.2), this fringe width on the observer plane is of order $\sim \lambda/\alpha$ where λ is wavelength and α is deflection angle. On the other hand, in order to resolve the multiple image formed by gravitational lensing, the observer size must be larger than $\sim \lambda/\alpha$. Therefore it is impossible to observe both the interference pattern and the multiple image simultaneously, in analogy to the uncertainty principle in quantum mechanics.

§4. Application to specific lens models

In this section some examples of wave effects are discussed for the point mass lens model and the fold caustic.

4.1. *Point mass lens*[5) - 7), 10)]

Lensing by a normal star is described by this lens model. If M is the mass of the star then $U(r,\theta) = -GM/(r^2 - 2r_l r\cos\theta + r_l^2)^{1/2}$. Substituting this into Eq. (2·6) and using $\theta \ll 1$ the lens potential is $\hat{\psi}(\theta_l) = 4GM\ln\theta_l + \mathrm{const}$, where the constant term is unimportant since it only affects the overall phase factor of F. Choosing the characteristic angle scale θ_* to be the Einstein angle,

$$\theta_* = \left(\frac{4GMr_{l0}}{r_l r_0}\right)^{1/2} \sim 3\mu\mathrm{arcsec}\left(\frac{M}{M_\odot}\right)^{1/2}\left(\frac{r_l r_0/r_{l0}}{\mathrm{Gpc}}\right)^{-1/2}, \tag{4·1}$$

one obtains from Eqs. (2·9) and (2·10)

$$\psi(x) = \ln x\,, \quad w = 4GM\omega \sim 10^5 (M/M_\odot)(\nu/\mathrm{GHz})\,. \tag{4·2}$$

Note that $w \sim$ (gravitational radius of lens mass)/(wavelength) is a large number in usual astrophysical situations. In any case Eq. (2·12) is integrated analytically to give the amplification factor for monochromatic waves from a point source as[19)]

$$|F|^2 = \frac{\pi w}{1-e^{-\pi w}}\left|{}_1F_1\left(\tfrac{1}{2}iw, 1; \tfrac{1}{2}iwy^2\right)\right|^2, \tag{4·3}$$

where ${}_1F_1$ is the confluent hypergeometric function. In the long wavelength limit $w \to 0$, Eq. (4·3) is identically unity and hence amplification does not occur (i.e., waves "ignore" the lens). When $w \gg 1$, Eq. (4·3) is evaluated accurately by the semi-classical approximation valid for $y \gg 1/w$, and by $|F|^2 \simeq \pi w\, J_0^2(wy)$ valid for $y \ll w^{-1/2}$ (cf. Eq. (3·8)). The time delay function $T(\boldsymbol{x}) = \frac{1}{2}|\boldsymbol{x}-\boldsymbol{y}|^2 - \ln x$ has two stationary points, a minimum point $\boldsymbol{x}_1$ and a saddle point $\boldsymbol{x}_2$ (cf. Eq. (3·6)):

$$\boldsymbol{x}_{1,2} = \frac{\boldsymbol{y}}{2y}(y \pm \sqrt{y^2+4})\,, \quad \mu_{1,2} = \frac{1}{2} \pm \frac{y^2+2}{2y\sqrt{y^2+4}}\,, \tag{4·4}$$

$$T_{12} = \frac{1}{2} y \sqrt{y^2+4} + \ln \frac{\sqrt{y^2+4}+y}{\sqrt{y^2+4}-y}\,, \quad n_{12} = \frac{1}{2}\,. \tag{4·5}$$

Thus the semi-classical approximation of Eq. (4·3) is

$$|F|^2 = \frac{y^2+2+2\sin(w\,T_{12})}{y\sqrt{y^2+4}}\,. \tag{4·6}$$

When $y \ll 1$, Eq. (4·6) becomes $|F|^2 \simeq [1+\sin(2wy)]/y$. The sine term represents the interference between the two images producing circular fringes. The critical curve (cf. §3.2) is a circle of radius $x = 1$ (Einstein ring) on the lens plane, and the caustic is the single point $y = 0$ on the source plane where Eq. (4·6) diverges. As discussed in §3.2, however, Eq. (4·6) is invalid when the time delay between the two images, $T_{12} \simeq 2y$, is comparable to or smaller than the wave period $1/w$. In reality the amplification $|F|^2$ at $y = 0$ remains finite due to the diffraction and has a maximum value $\pi w/(1-e^{-\pi w})$ according to Eq. (4·3). The maximum amplification $\simeq \pi w$ is of order $\sim$ (gravitational radius)/(wavelength) when this ratio is much larger than unity.

Next we consider an extended source. As discussed in §3.3, the interference term $\sin(w\,T_{12})$ is erased when the source size is much larger than the width between interference fringes. This width is π/w at $y \ll 1$ for waves of frequency w, and becomes narrower for larger y. Thus the conclusion is that if the source size is much larger than $r_0\,\theta_*(\pi/w) =$

$$\frac{\pi}{\omega}\sqrt{\frac{r_0 r_{l0}}{4GMr_l}} \sim 10^7 \text{ km} \left(\frac{\nu}{\text{GHz}}\right)^{-1} \left(\frac{M}{M_\odot}\right)^{-1/2} \left(\frac{r_0 r_{l0}/r_l}{\text{Gpc}}\right)^{1/2}\,, \tag{4·7}$$

then the geometric optics approximation, $|F|^2 = (y^2+2)/(y\sqrt{y^2+4})$, is satisfactory. In Fig. 3 is plotted the amplification factor for a circular source of radius Δy whose center is on the caustic $y = 0$. Here uniform brightness and a monochromatic spectrum are assumed: $I(w', y') \propto \mathrm{H}(\Delta y - y')\,\delta(w'-w)$ where H is the step function. The solid curves are the full results $C(0) = 2\int_0^{\Delta y} dy\, y|F|^2/(\Delta y)^2$ (cf. Eq. (2·15)) with $w = 10^2$, 10^3 and 10^4 from bottom to top, and the dashed curve is the geometric optics approximation $C_0(0) = [1+(2/\Delta y)^2]^{1/2}$ (cf. Eq. (3·10)). This figure reveals that the diffraction as well as the interference is unimportant when the source size is much larger than Eq. (4·7).

As for non-monochromatic waves, the interference terms $C_1(0)$ and $C_2(0)$ (Eq. (3·11)) vanish even for a point source, when the coherence time $1/\Delta w$ is much smaller than the time delay T_{12} (cf. §3.3). For example, for a point source and a uniform spectrum $I(w', y') \propto \delta^2(\boldsymbol{y}' - \boldsymbol{y})\,\mathrm{H}(\Delta w - |w'-w|)$, $C(0)$ becomes Eq. (4·6) with the sine term multiplied by $j_0(\Delta w\,T_{12})$ (where $j_0(x) = \sin x/x$). Therefore the wave effects are unimportant when $\Delta w\,T_{12} \gg 1$ as regards $C(0)$. However, if one observes not only $C(0)$ but also $C(\tau)$ with variable τ, then it is possible to detect C_1 or C_2 regardless of Δw. In Fig. 4 $|C_1(\tau)|$ is plotted for a circular source located at $y = 1$. A uniform brightness and a Gaussian spectrum, $\tilde{I}(\tau, \boldsymbol{y}') \propto \mathrm{H}(\Delta y - |\boldsymbol{y}' - \boldsymbol{y}|)\,\exp[-(\Delta w\,\tau/2)^2]$, are assumed. Since C_0 and C_2 are negligible

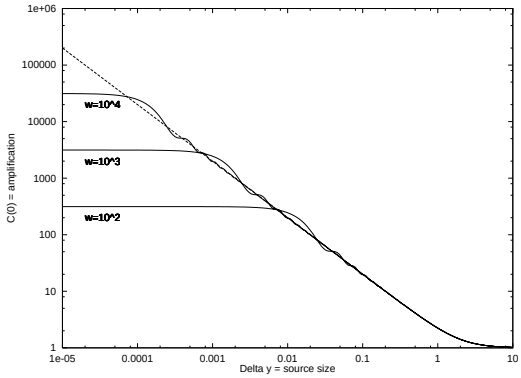

Fig. 3. Amplification factor by a point mass lens for a circular source of radius Δy with its center located at $y = 0$. A uniform brightness and monochromatic spectrum are assumed. The solid curves are the full results $C(0)$ with frequencies $w = 10^2$, 10^3 and 10^4 from bottom to top. The dashed curve is the geometric optics approximation $C_0(0)$. The abscissa at $\Delta y = \pi/w$ corresponds to Eq. (4·7). When $\Delta y \to 0$, $C(0)$ tends to πw.

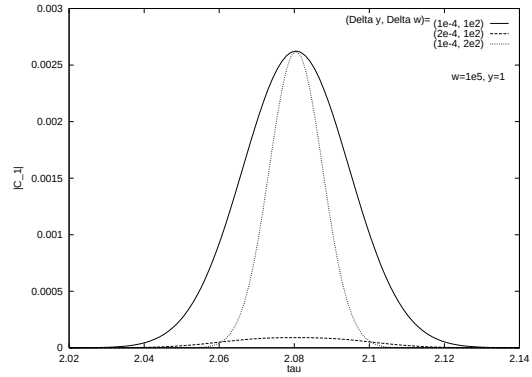

Fig. 4. Interference term $|C_1(\tau)|$ in the correlation function for a circular source located at $y = 1$. Solid, dashed and dotted curves correspond to $(\Delta y, \Delta w) = (10^{-4}, 10^2)$, $(2 \cdot 10^{-4}, 10^2)$ and $(10^{-4}, 2 \cdot 10^2)$, respectively. A uniform brightness and Gaussian spectrum are assumed. The central frequency of the spectrum is $w = 10^5$.

in the plotted range of τ (thanks to the Gaussian filter), this plot represents the amplitude of the full correlation function $C(\tau)$. It is seen that $C(\tau)$ peaks around $\tau \sim T_{12}|_{y=1} \sim 2.08$ and that the peak height decreases with the source size Δy. Searching for this peak in the $C(\tau)$ curve would be a method of detecting the tiny interference term as evidence of gravitational lensing, even if the source size is much larger than Eq. (4·7).[5),6)]

4.2. *Diffraction near fold singularity*[10),13),20)]

A fold caustic is a one-dimensional curve on the $\boldsymbol{y}$-plane characterized by $\mathrm{rank}[\partial_a \partial_b T(\boldsymbol{x})] = 1$.[21)] Examples of fold (and cusp) caustic are found in the elliptical lens model for the galaxy,[22)] a point mass lens perturbed by external shear[23)] (cf. Fig. 6), a collection of point mass lenses,[24)] etc.

A point source of monochromatic waves is located at $\boldsymbol{y}$ close to the fold caustic. Let $\boldsymbol{y}_0$ be a point on the caustic closest to the source, and let $\boldsymbol{x}_0$ be the corresponding point on the critical curve satisfying $\boldsymbol{y}_0 = \boldsymbol{x}_0 - \boldsymbol{\partial}\psi(\boldsymbol{x}_0)$ (cf. Eq. (3·1)). In computing the diffraction integral (Eq. (2·11)), we neglect possible contributions from stationary points of $T(\boldsymbol{x})$ far from $\boldsymbol{x}_0$, since the integral near $\boldsymbol{x}_0$ should be the dominant contribution to the wave amplitude. We set Cartesian coordinates $\boldsymbol{x} = (x_1, x_2)$ and $\boldsymbol{y} = (y_1, y_2)$ such that $\partial_1^2 T(\boldsymbol{x}_0)$ and $\partial_1 \partial_2 T(\boldsymbol{x}_0)$ vanish using $\mathrm{rank}[\partial_a \partial_b T(\boldsymbol{x}_0)] = 1$. (Then the tangent vector to the caustic at $\boldsymbol{y}_0$ is parallel to the y_2-axis.) The time delay $T(\boldsymbol{x}) = \frac{1}{2}|\boldsymbol{x} - \boldsymbol{y}|^2 - \psi(\boldsymbol{x})$ is expanded around $\boldsymbol{x}_0$ as

$$T(\boldsymbol{x}) = T(\boldsymbol{x}_0) - \bar{\boldsymbol{x}} \cdot \bar{\boldsymbol{y}} + \frac{\bar{x}_2^2}{2p} + \frac{1}{6} \sum_{abc} \bar{x}_a \bar{x}_b \bar{x}_c \, \partial_a \partial_b \partial_c T(\boldsymbol{x}_0) + \cdots, \qquad (4\cdot8)$$

where $\bar{\boldsymbol{x}} = \boldsymbol{x} - \boldsymbol{x}_0$, $\bar{\boldsymbol{y}} = \boldsymbol{y} - \boldsymbol{y}_0$, and $p = 1/[\partial_2^2 T(\boldsymbol{x}_0)] = 1/[1 - \partial_2^2 \psi(\boldsymbol{x}_0)]$ is assumed to be of order unity. Due to the absence of a $\bar{x}_1^2$ term, the third order terms contribute to the diffraction integral in the $\bar{x}_1$-direction. The fold caustic further requires the non-vanishing of $\partial_1^3 T(\boldsymbol{x}_0)$ (which guarantees the non-vanishing of the tangent vector to the fold caustic at $\boldsymbol{y}_0$), and we choose θ_* (Eq. (2·10)) so that $q = 2/[\partial_1^3 T(\boldsymbol{x}_0)] = -2/[\partial_1^3 \psi(\boldsymbol{x}_0)]$ is of order unity. In the case of a point mass lens with external shear [23] or a collection of point mass lenses, [24] the choice of θ_* in Eq. (4·1) makes q order unity. Without loss of generality q is taken to be positive by inversion of coordinates ($\bar{\boldsymbol{x}} \leftrightarrow -\bar{\boldsymbol{x}}$ and $\bar{\boldsymbol{y}} \leftrightarrow -\bar{\boldsymbol{y}}$). We omit all the third order terms except the $\bar{x}_1^3$ term, since these terms ($\bar{x}_1^2 \bar{x}_2$, $\bar{x}_1 \bar{x}_2^2$ and $\bar{x}_2^3$) do not contribute to the wave amplitude (Eq. (2·11)) in the short wavelength limit $w \to \infty$. Without these terms the caustic and critical lines coincide with the $\bar{y}_2$-axis and $\bar{x}_2$-axis, respectively, and the amplification factor of the wave intensity is

$$|F|^2 = \frac{w^2}{4\pi^2} \left| \int d\bar{x}_1 \int d\bar{x}_2 \, \exp\left[iw \left(-\bar{x}_1 \bar{y}_1 - \bar{x}_2 \bar{y}_2 + \frac{\bar{x}_2^2}{2p} + \frac{\bar{x}_1^3}{3q} \right) \right] \right|^2, \tag{4·9}$$

$$= 2\pi \mu_* \, \mathrm{Ai}^2(-\bar{y}_1/Y) \,, \tag{4·10}$$

where $\mu_* = |p| \, (q^2 w)^{1/3}$, $Y = (qw^2)^{-1/3}$ and Ai is the Airy function.

When $|\bar{y}_1|^{3/2} \gg 1/w$, Eq. (4·10) is evaluated accurately by the semi-classical approximation. On the positive side $\bar{y}_1 > 0$ of the caustic, the phase of Eq. (4·9) has two stationary points (cf. Eq. (3·6)):

$$\bar{\boldsymbol{x}}_\pm = (\pm\sqrt{q\,\bar{y}_1}, p\,\bar{y}_2) \,, \quad \mu_\pm = \pm \tfrac{1}{2} p (q/\bar{y}_1)^{1/2} \,, \tag{4·11}$$

$$\Delta T = T(\boldsymbol{x}_-) - T(\boldsymbol{x}_+) = \tfrac{4}{3} (q\,\bar{y}_1^3)^{1/2} \,, \quad n_- - n_+ = 1/2 \,. \tag{4·12}$$

On the negative side $\bar{y}_1 < 0$, there are not real stationary points but are imaginary ones. Hence the semi-classical approximation of Eq. (4·10) is

$$|F|^2 = |p| (q/|\bar{y}_1|)^{1/2} \times \begin{cases} 1 + \sin(w\,\Delta T), & (\bar{y}_1 > 0) \\ \frac{1}{2} \exp(-w|\Delta T|). & (\bar{y}_1 < 0) \end{cases} \tag{4·13}$$

In Fig. 5(a) $|F|^2$ is plotted as a functioin of $\bar{y}_1$. As a point source crosses the caustic line $\bar{y}_1 = 0$ from its positive side to its negative side with fixed $\bar{y}_2$, one observes two images of equal brightness which approach each other, merge on the critical line $\bar{x}_1 = 0$, and finally disappear. The sine term expresses the interference between these two images, and according to Eq. (4·13) the wave intensity diverges at the instant of merging. As discussed in §3.2, however, the semi-classical approximation is invalid when the source is so close to the caustic that the time delay between the two images, $|\Delta T| \sim |\bar{y}_1|^{3/2}$, is comparable to or smaller than the wave period $1/w$. In reality, according to Eq. (4·10), the amplification factor $|F|^2$ has a finite value of $0.792\mu_*$ at $\bar{y}_1 = 0$, and a maximum value of $1.803\mu_*$ at $\bar{y}_1 = 1.019Y$. Despite the absence of images on the negative side of the caustic, the wave intensity on that side is not exactly zero but decreases exponentially with $|\bar{y}_1|$, due to the diffraction of waves from the positive side.

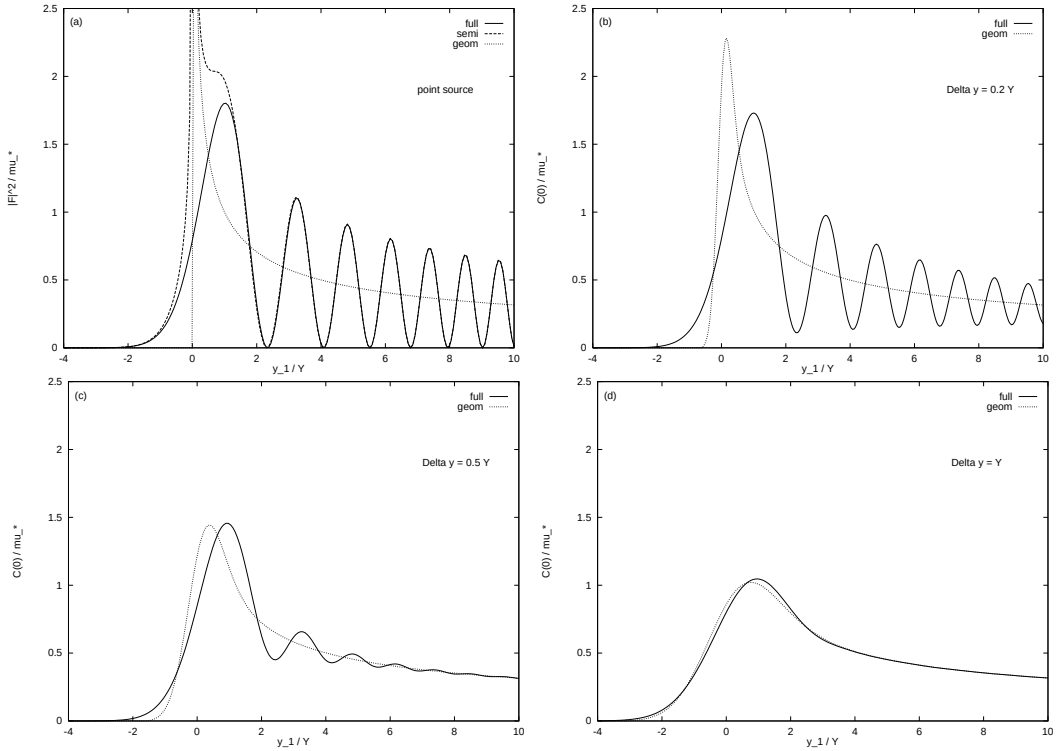

Fig. 5. (a) Amplification factor near the fold caustic for a point source. Solid, dashed and dotted curves represent the full result in Eq. (4·10), the semi-classical approximation in Eq. (4·13), and the geometric optics approximation (Eq. [4·13) without the sin and exp terms), respectively. (b)–(d) Amplification factor for extended sources of size $\Delta y = 0.2Y$, $0.5Y$ and Y. The Gaussian brightness profile and monochromatic spectrum are assumed. Solid and dotted curves show the full result $C(0)$ and the geometric optics approximation $C_0(0)$, respectively.

The width between interference fringes is $\sim Y \sim w^{-2/3}$ at small $\bar{y}_1$ and becomes narrower for larger $\bar{y}_1$. Therefore, from the discussion in §3.3, if the size of an extended source is much larger than $r_0\,\theta_*/w^{2/3} \sim$

$$\left[\frac{r_0 r_{l0}/r_l}{(4GM\omega^4)^{1/3}}\right]^{1/2} \sim 2\cdot 10^8 \text{ km} \left(\frac{\nu}{\text{GHz}}\right)^{-2/3} \left(\frac{M}{M_\odot}\right)^{-1/6} \left(\frac{r_0 r_{l0}/r_l}{\text{Gpc}}\right)^{1/2}, \quad (4\cdot14)$$

then the geometric optics approximation, $|F|^2 = \mu_*(\bar{y}_1/Y)^{-1/2}\mathrm{H}(\bar{y}_1)$, is satisfactory (where H is the step function). In Fig. 5(b)–(d) is plotted the amplification factor for extended sources of various sizes $\Delta y = 0.2Y$, $0.5Y$ and Y. A Gaussian brightness profile and monochromatic spectrum, $I(w', \boldsymbol{y}') \propto \exp[-(|\boldsymbol{y}' - \boldsymbol{y}|/\Delta y)^2/2]\,\delta(w' - w)$, are assumed. (Figure 5 is independent of the value of w.) Solid and dotted curves represent the full result $C(0)$ (cf. Eq. (2·15)) and the geometric optics approximation $C_0(0)$ (cf. Eq. (3·10)), respectively. This figure verifies the above statement that both the diffraction and the interference are insignificant when the source size is much larger than Eq. (4·14).

§5. Numerical method of performing the diffraction integral

If the semi-classical approximation Eq. (3·6) is invalid because of diffraction, the diffraction integral Eq. (2·11) need to be evaluated exactly. In this section a numerical method [14] of evaluating Eq. (2·11) for a general lens potential $\psi(\boldsymbol{x})$ is presented. First, one calculates the following quantity for a fixed $\boldsymbol{y}$:

$$\tilde{F}(\tau) = \int^{\tau} d\tau' \int_{-\infty}^{\infty} \frac{dw}{2\pi}\, e^{-iw\tau'} F(w) = \frac{1}{2\pi} \int d^2\boldsymbol{x}\ \delta[T(\boldsymbol{x}) - \tau]\,. \tag{5·1}$$

Note that $d\tilde{F}/d\tau$ represents the lensed waveform for a delta-function pulse emitted from a point source, and $\tilde{F}$ is defined up to an additive constant. Given a value of τ, closed contour curves are determined by $\tau = T(\boldsymbol{x})$ on the $\boldsymbol{x}$-plane (e.g. Fig. 7). Equation (5·1) is written in terms of line integrals along these contours C as

$$\tilde{F}(\tau) = \frac{1}{2\pi} \sum_{C} \oint_{C} du\,, \tag{5·2}$$

where the summation is taken over the number of closed contours, and

$$\frac{dx_1}{du} = -(x_2 - y_2) + \partial_2\psi(\boldsymbol{x})\,, \quad \frac{dx_2}{du} = x_1 - y_1 - \partial_1\psi(\boldsymbol{x}) \tag{5·3}$$

in Cartesian coordinates $\boldsymbol{x} = (x_1, x_2)$ and $\boldsymbol{y} = (y_1, y_2)$, so that $d\boldsymbol{x}/du$ is tangent to the contours with length $|\boldsymbol{\partial} T|$. The differential equations (5·3) are integrated numerically using, e.g., the Runge-Kutta method to evaluate Eq. (5·2). The integration is started with $u = 0$ from a point on the contour and is continued with positive steps of u until $\boldsymbol{x}$ comes back to the starting point. Computing $\tilde{F}$ in this way, one obtains F from the Fourier transform of $d\tilde{F}/d\tau$:

$$F(w) = \int_{-\infty}^{\infty} d\tau\, e^{iw\tau} \frac{d}{d\tau} \tilde{F}(\tau)\,. \tag{5·4}$$

Since $d\tilde{F}/d\tau$ is the lensed waveform for a delta-function pulse, $\tilde{F}(\tau)$ should exhibit a discontinuous or diverging behavior when τ is equal to the arrival time T_j of the pulse (T_j is the value of $T(\boldsymbol{x})$ at its stationary point $\boldsymbol{x}_j$; cf. Eq. (3·4)). Also the number of closed contours changes by one when $\tau = T_j$, while it is zero when $\tau \to -\infty$. Assuming that the source is not exactly on a caustic, one can show [14] that $\tilde{F}(\tau)$ at $\tau = T_j$ suffers a discontinuous increase (decrease) by an amount ${\mu_j}^{1/2}$ when $\boldsymbol{x}_j$ is a minimum (maximum) point, and that it diverges as $\pi^{-1}|\mu_j|^{1/2} \ln|\tau - T_j|^{-1}$ when $\boldsymbol{x}_j$ is a saddle point (e.g. Fig. 8). Let us write these discontinuous contributions to $\tilde{F}$ as

$$\tilde{F}_{\rm c}(\tau) = \sum_{n_j=0,1} (-1)^{n_j} {\mu_j}^{1/2}\, \mathrm{H}(\tau - T_j) - \frac{1}{\pi} \sum_{n_j=1/2} |\mu_j|^{1/2}\, \ln|\tau - T_j|\,, \tag{5·5}$$

where H is the step function. Note that $F_{\rm c} = \int_{-\infty}^{\infty} d\tau\, e^{iw\tau} d\tilde{F}_{\rm c}/d\tau$ (cf. Eq. (5·4)) is equal to[*] the semi-classical approximation of F, given in Eq. (3·4). Hence Eq. (5·5) is

[*] using $\int_{-\infty}^{\infty} dx\, e^{ix}/x = i\pi$

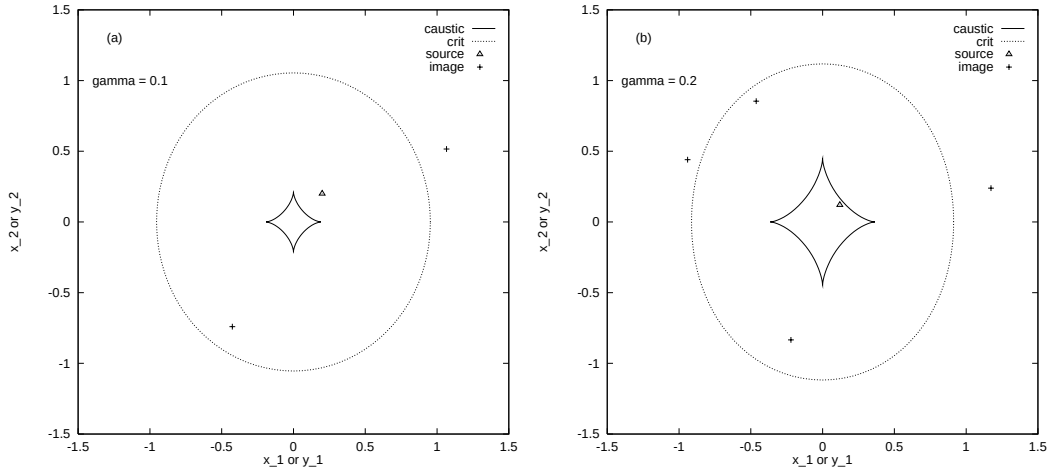

Fig. 6. Point mass lens perturbed by external shear. The critical curve (dotted curve) is expressed as $x(\varphi) = [\gamma \cos 2\varphi + (1 - \gamma^2 \sin^2 2\varphi)^{1/2}]^{-1/2}$ in circular coordinates. The caustic (solid curve) is the diamond-shaped curve with four cusps. The image positions (crosses) are the stationary points of Eq. (5·7). The number of images is two or four depending on whether the source (triangle) is inside or outside the caustic. (a) $\gamma = 0.1$. (b) $\gamma = 0.2$.

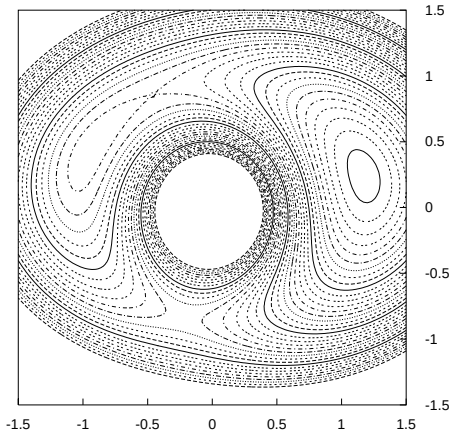

Fig. 7. Contours of the time delay function $T(\boldsymbol{x})$ (Eq. (5·7)) in the case of Fig. 6(b).

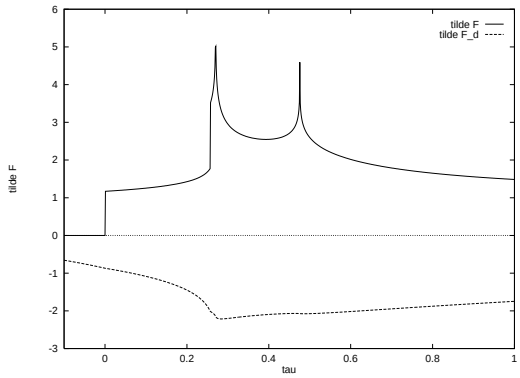

Fig. 8. Solid: lensed waveform for a delta function pulse $\tilde{F}(\tau)$ (Eq. (5·2)) in the case of Fig. 6(b). Dashed: diffraction part of $\tilde{F}(\tau)$ as obtained by subtracting Eq. (5.5) from Eq. (5·2).

regarded as the semi-classical part of $\tilde{F}$. The remaining diffraction part $F_{\rm d} = F - F_{\rm c}$ is obtained by applying $\tilde{F}_{\rm d} = \tilde{F} - \tilde{F}_{\rm c}$ to Eq. (5·4):

$$F_{\rm d}(w) = e^{iw\tau} \tilde{F}_{\rm d}(\tau)|_{\tau_1}^{\tau_2} - iw \int_{\tau_1}^{\tau_2} d\tau \, e^{iw\tau} \tilde{F}_{\rm d}(\tau) \, . \qquad (5\cdot6)$$

The integral range (τ_1, τ_2) contains all the T_j but is kept finite in actual computations, because $\tilde{F}_{\rm d}(\tau)$ is logarithmically divergent due to the second term of Eq. (5·5).*)

*) The first term of Eq. (5·6) is absent in the original reference;[14] there it is not noted that $\tilde{F}$ is defined up to an additive constant.

Thus after removing the discontinuous part $\tilde{F}_{\rm c}$ from $\tilde{F}$, it is possible to calculate reliably the second term of Eq. (5·6) for a continuous function $\tilde{F}_{\rm d}$ using, e.g., the Fast-Fourier-Transform. To summarize, i) compute Eq. (5·2) solving the differential Eqs. (5·3), ii) subtract Eq. (5·5) from Eq. (5·2) and insert the result into Eq. (5·6), and iii) add Eqs. (3·4) and (5·6) to obtain Eq. (2·11).

We apply the above algorithm to the lens model of a point mass plus external shear: [23]

$$T(\boldsymbol{x}) = \tfrac{1}{2}|\boldsymbol{x}-\boldsymbol{y}|^2 - \ln|\boldsymbol{x}| - \tfrac{1}{2}\gamma(x_1^2 - x_2^2)\,. \tag{5·7}$$

The shear parameter γ (we assume $0<\gamma<1$) accounts for the tidal gravitational field from an external perturber in the x_2 direction. For example, the perturber may be another point mass lens at $\boldsymbol{x}=(0,\gamma^{-1/2})$, or may be the galactic potential that contains the main lens. Figures 6(a) and 6(b) display the lensing configurations for which we calculate the wave amplitude, with $\gamma=0.1$ and 0.2. In Fig. 6(b) the source is near the caustic so the two images near the critical curve have large magnifications. Figure 7 displays the contours of Eq. (5·7) for the configuration in Fig. 6(b). The image positions in Fig. 6(b) coincide with the stationary points of $T(\boldsymbol{x})$ in Fig. 7. It is seen that the number of closed contours (C in Eq. (5·2)) changes as 0,1,2,1.2 when τ moves from $-\infty$ to ∞. In Fig. 8 $\tilde{F}(\tau)$ is plotted in the case of Fig. 6(b) as computed by Eqs. (5·2) and (5·3). Corresponding to the two minima and two saddle points in Fig. 7, the solid curve has two discontinuities and two logarithmic spikes. Subtracting the semi-classical contribution (Eq. (5·5)) from the solid curve yields the diffraction part $\tilde{F}_{\rm d}$ plotted as the dashed curve. Finally Figs. 9(a) and 9(b) show the amplification factor $|F|^2$ as a function of the frequency w for a point source in Figs. 6(a) and 6(b). The dotted, dashed and solid curves correspond to the semi-classical approximation $|F_{\rm c}|^2$ (Eq. (3·6)), the contribution from the diffraction $|F_{\rm d}|^2$ (Eq. (5·6)) which is noticeable at low frequencies, and the full result $|F_{\rm c}+F_{\rm d}|^2$, respectively. In Fig. 9(a) $|F|^2$ simply oscillates with period $2\pi/\Delta T$, corresponding to the time delay $\Delta T\sim 0.6$ between the two images in Fig. 6(a). On the other hand,

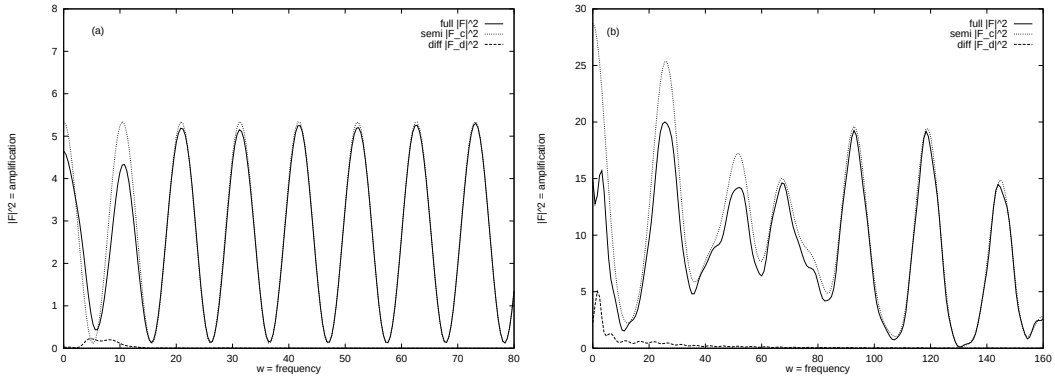

Fig. 9. Amplification factor $|F|^2$ versus frequency w for a point source in Figs.6(a) and (b). Solid: full result (Eq. (2·11)). Dotted: semi-classical approximation (Eq. (3·6)). Dashed: contribution from diffraction (Eq. (5·6)).

the spectrum in Fig. 9(b) is rather complex because $|F_c|^2$ has six interference terms out of the four images in Fig. 6(b). Also in Fig. 9(b) the diffraction is noticeable up to high frequency $w \sim 80$ because of the small time delay $\Delta T \sim 0.013$ between the two bright images near the critical curve. This figure exemplifies the conclusion in §3.2 that the diffraction effect is significant when a point source is located near a caustic.

§6. Summary

From the general discussion given in §§2 and 3 and from the examples given in §§4 and 5, we summarize this paper as follows; (1) If the wavelength is comparable to or larger than the gravitational radius of the lens mass, then diffraction is so effective that large amplification cannot occur; (2) If the time delay between the lensed images is comparable to or smaller than the wave period, then diffraction is important (cf. Fig. 9). This happens when a source is located near a caustic. The condition here also implies the condition in (1), since the time delay is of order $\sim$ (gravitational radius)/(light velocity); (3) If the source size is much larger than the width between interference fringes $\sim$ (wavelength)/(deflection angle), then both diffraction and interference are unimportant so that geometric optics is valid (cf. Fig. 3); (4) If the coherence time (band width)$^{-1}$ is much smaller than time delay between images, then geometric optics is valid; (5) If one observes the correlation function $C(\tau)$ (Eq. (2·14)) with variable τ, then it is possible to detect interference terms regardless of the condition in (4) (cf. Fig. 4).

Wave effects in gravitational lensing have not yet been observed. We do not know of compact radio sources much smaller than Eq. (4·7) at cosmological distances.[25] The sources of gamma-ray bursts should be sufficiently compact,[26] but "femtolenses" of 10^{-13}–$10^{-16} M_\odot$ are hypothetical objects.[27] Gravitational waves emitted from binary stars are likely to exhibit the effects considered in this paper if lensed,[16] but the detection rate of these waves for planned detectors is uncertain.[28] In conclusion the geometric optics approximation of gravitational lensing is valid in all the observed situations.

Acknowledgements

The first author thanks M. Sasaki and A. Hosoya for discussions on §§2.1 and 3.3, respectively. We are grateful to K. Tomita for the opportunity to publish this review.

Appendix

The first $l-1$ integrals from $j = 1$ to $l-1$ in Eq. (2·7) are the Gaussian integrals and give unity, so $j = 1$ under $\prod$ and $\sum$ can be replaced by $j = l$. We use the identity

(proven by induction)

$$\sum_{j=l}^{N-1} r_j r_{j+1} |\boldsymbol{\theta}_j - \boldsymbol{\theta}_{j+1}|^2 = \epsilon \frac{r_l r_0}{r_{l0}} |\boldsymbol{\theta}_l - \boldsymbol{\theta}_0|^2 + \sum_{j=l+1}^{N-1} r_j^2 \frac{r_{l,j+1}}{r_{lj}} |\boldsymbol{\theta}_j - \boldsymbol{u}_{lj}|^2 , \qquad \text{(A·1)}$$

where $r_0 = r_N$, $r_{lj} = r_j - r_l$ and $\boldsymbol{u}_{lj} = [r_l \boldsymbol{\theta}_l + (j-l) r_{j+1} \boldsymbol{\theta}_{j+1}]/(j\, r_{l,j+1})$. Thus Eq. (2·7) becomes

$$F(\vec{r}_0) = \int \frac{d^2\boldsymbol{\theta}_l}{A_l} \exp\Big\{ i\omega \Big[\frac{r_l r_0}{2 r_{l0}} |\boldsymbol{\theta}_l - \boldsymbol{\theta}_0|^2 - \hat{\psi}(\boldsymbol{\theta}_l) \Big] \Big\} \times \left[\prod_{j=l+1}^{N-1} \int \frac{d^2\boldsymbol{\theta}_j}{A_j} \right] \exp\left[\frac{i\omega}{2\epsilon} \sum_{j=l+1}^{N-1} r_j^2 \frac{r_{l,j+1}}{r_{lj}} |\boldsymbol{\theta}_j - \boldsymbol{u}_{lj}|^2 \right] . \qquad \text{(A·2)}$$

Evaluating the Gaussian integrals in the second line of this equation results in

$$\text{2nd line of (A·2)} = \prod_{j=l+1}^{N-1} \frac{1}{A_j} \frac{2\pi i \epsilon}{\omega\, r_j^2} \frac{r_{lj}}{r_{l,j+1}} = \prod_{j=l+1}^{N-1} \frac{r_{j+1}}{r_j} \frac{r_{lj}}{r_{l,j+1}} = \frac{r_0}{r_{l+1}} \frac{\epsilon}{r_{l0}} . \qquad \text{(A·3)}$$

Substituting this into Eq. (A·2) yields Eq. (2·8).

References

1) H. C. Ohanian, Int. J. Theor. Phys. **9** (1974), 425.
2) P. V. Bliokh and A. A. Minakov, Astrophys. Space Sci. **34** (1975), L7.
3) R. J. Bontz and M. P. Hogan, Astrophys. Space Sci. **78** (1981), 199.
4) A. V. Mandzhos, Sov. Astron. Lett. **7** (1982), 213.
5) P. Schneider and J. Schmidt-Bergk, Astron. Astrophys. **148** (1985), 369.
6) S. Deguchi and W. D. Watson, Phys. Rev. **D34** (1986), 1708.
7) S. Deguchi and W. D. Watson, Astrophys. J. **307** (1986), 30.
8) S. Deguchi and W. D. Watson, Astrophys. J. **315** (1987), 440.
9) J. J. Goodman, R. W. Romani, R. D. Blandford and R. Narayan, Mon. Not. R. Astron. Soc. **229** (1987), 73.
10) P. Schneider, J. Ehlers and E. E. Falco, *Gravitational Lenses,* (Springer-Verlag, Berlin, 1992), §7.
11) A. Gould, Astrophys. J. **386** (1992), L5.
12) K. Z. Stanek, B. Paczyński and J. Goodman, Astrophys. J. **413** (1993), L7.
13) M. Jaroszyński and B. Paczyński, Astrophys. J. **455** (1995), 443.
14) A. Ulmer and J. Goodman, Astrophys. J. **442** (1995), 67.
15) A. Gould and B. S. Gaudi, Astrophys. J. **486** (1997), 687.
16) T. T. Nakamura, Phys. Rev. Lett. **80** (1998), 1138.
17) R. P. Feynmann, Rev. Mod. Phys. **20** (1948), 267.
18) C. W. Misner, K. S. Thorne and J. A. Wheeler, *Gravitation* (Freeman, San Francisco, 1973), §§22.5 and 22.6.
19) I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, series and products* (Academic Press, New York, 1980), Eq. (6.631.1).
20) M. V. Berry and C. Upstill, Prog. Optics **18** (1980), 257.
21) R. Blandford and R. Narayan, Astrophys. J. **310** (1986), 568.
22) R. D. Blandford and C. S. Kochanek, Astrophys. J. **321** (1987), 658.
23) K. Chang and S. Refsdal, Astron. Astrophys. **132** (1984), 168.
24) e.g., J. Wambsganss, B. Paczyński and P. Schneider, Astrophys. J. **358** (1990), L33.
25) e.g., L. Kedziora-Chudczer, K. L. Jauncey, M. H. Wieringa, M. A. Walker, G. D. Nicolson, J. E. Reynolds and A. K. Tzioumis, Astrophys. J. **490** (1997), L9.
26) e.g., T. Piran, Gen. Rel. Grav. **28** (1996), 1421.
27) E. W. Kolb and I. I. Tkachev, Astrophys. J. **460** (1996), L25.
28) E. S. Phinny, Astrophys. J. **380** (1991), L17.