On the theory of Complex Rays

S. J. Chapman
J. M. H. Lawry
J. R. Ockendon
OCIAM
Mathematical Institute
24–29 St. Giles'
Oxford OX1 3LB

R. H. Tew
Department of Theoretical Mechanics
University of Nottingham
University Park
Nottingham NG7 2RD

October 15, 1999

Abstract

The article surveys the application of complex ray theory to the scalar Helmholtz equation in two dimensions.

The first objective is to motivate a framework within which complex rays may be used to make predictions about wavefields in a wide variety of geometrical configurations. A crucial ingredient in this framework is the rôle played by Stokes' phenomenon in determining the regions of existence of complex rays. The identification of the Stokes surfaces emerges as a key step in the approximation procedure, and this leads to the consideration of the many characterisations of Stokes surfaces, including the adaptation and application of recent developments in exponential asymptotics to the complex WKB expansion of these wavefields.

Examples are given for several cases of physical importance.

1 Introduction

The objective of this paper is to lay down a systematic complexified theory of monochromatic high-frequency wave propagation. The computational and analytical usefulness of such a theory for real rays has been very well developed, as described in [24]. However, it is usually only the real solutions of the eikonal equation that are taken into account, with complex solutions being tacitly ignored. Moreover, those few complexifications of real ray theory that have been proposed [25, 26, 43] sometimes differ in their methodologies.

We will see that complex rays may be generated even in problems in which the incident field is composed of only real rays, and thus to describe fully the solution to any high-frequency wave propagation problem we should keep in mind the possibility of complex ray contributions. We draw particular attention to two types of problem for which the theory of complex rays is essential if we are to find even the lowest-order approximation to the rayfield. The first concerns situations in which there are no real rays in some sections of space, so that the ray approximation to the solution consists solely of complex rays. An example of such a situation is the circular caustic which we introduce in §2.2 and return to throughout the paper, in which

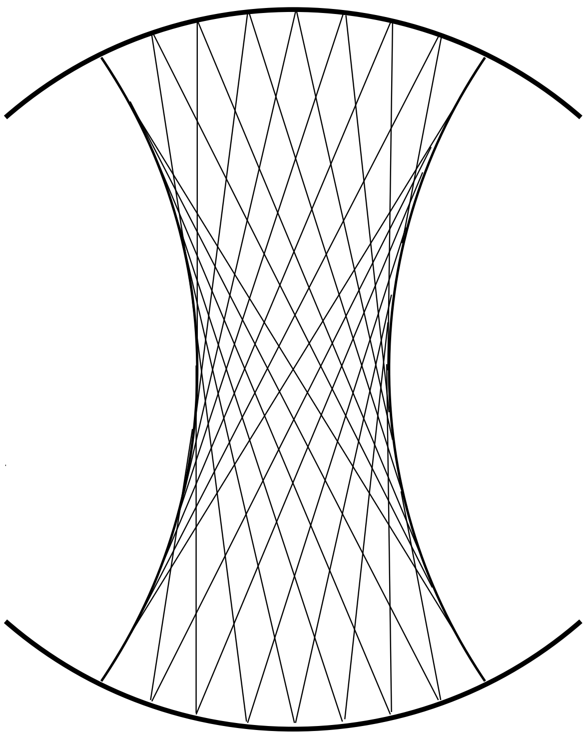


Figure 1: A trapped bouncing-ball mode between two curved endplates.

there are no real rays on the dark side of the caustic. A second example concerns a trapped bouncing-ball mode between two curved endplates, as illustrated in Figure 1. Here all the real rays are confined between the plates, but complex rays will be diffracted from the endpoints of the plates, which will leak an exponentially small amount of energy to infinity.

The second type of problem concerns situations in which the incoming field itself comprises complex rays. Such a situation is typified by problems involving insonification by Gaussian beams, which we will see in §2.3 can be described by an isotropic source located at a complex point in space.

In many other problems, complex rays are generated but are subdominant to real rays, and in these cases we aim to use knowledge of complex ray theory to describe the solution to a greater accuracy. This third type of problem is typified by the problem of total internal reflection of rays from a point source near a plane boundary, which we introduce in §2.3 and treat by complex-ray methods in §6. In this problem the complex ray is exponentially subdominant to the real rays in one halfspace, but becomes of the same order near the interface. Thus the complex ray contribution is very significant there. The behaviour near this boundary could alternatively be discerned through a boundary layer analysis (as in [42]), but since there is no singularity in the ray solution when complex rays are included the boundary layer is really artificial, and the solution to the boundary layer problem is simply the local behaviour of the full ray solution including complex rays.

A systematic procedure for describing complex rays involves not only the WKB and matched asymptotic expansions that underpin [24] but also the more subtle theories of exponential asymptotics that have only recently appeared in the literature [37]. Hence the goal of the first half of this paper is both to exemplify situations where it is necessary to consider complex solutions of the eikonal equation and also to emphasise the pivotal role played by "Stokes surfaces" concerning the physical interpretation of these solutions. These surfaces are defined in §4 and are multidimensional generalisations of the Stokes lines that are encountered when holomorphic functions of one variable are expanded in the neighbourhood of an isolated essential singularity.

Suppose a multidimensional rayfield exhibits a singular behaviour, such as a caustic, that necessitates a local modification of the ray ansatz. The leading term in the far-field expansion of the local asymptotic approximation will, in general, exhibit such an essential singularity, and thus Stokes surfaces will always exist in the vicinity of caustics and other places where two of more rayfields have a ray, real or complex, in common. The principal attribute of the Stokes surfaces is that it is across them, and only across them, that subdominant solutions of the Helmholtz equation can be "switched on or off." This makes complex-ray tracing qualitatively different from real-ray tracing over and above the obvious need for good four- (or six-) dimensional visualisations; whereas for real rays we can associate a directivity to ensure uniqueness of a physically acceptable solution, for complex rays we must replace this with the idea that relevant wavefields must be switched on or off when the ray passes through the Stokes surface.

The plan of this article is as follows. Having used some simple examples to motivate our ideas in §2, we give an informal discussion of the problems encountered in visualising complex rays in §3 and this allows us to build up to the all-important §4 where Stokes surfaces are defined, characterised and identified in a simple example, namely the case of a circular caustic. In §5 we proceed to study the further geometrical complexity introduced by singularities in the exact solution to the Helmholtz equation, again by reference to a simple example. In §6 we apply the theory we have developed to some more complicated examples, beginning with half-planes, where the geometry allows us to check our predictions against Fourier integrals. Our first new result concerns the prediction of complex rays in certain regions of real space for a physically realisable problem concerning total internal reflection. More daunting is the illumination of a smooth convex scatterer by a plane wave where a highly intricate Stokes surface is revealed, but, despite the presence of creeping rays, no complex rays exist in real space. By extending our analysis to complex plane wave illumination, we can however make new predictions about a scattering problem in which complex rays exist in real space. Finally, in §7, we discuss the effects of non-analyticity of the boundaries in order to construct a complex "Geometrical Theory of Diffraction". This is not as hard as it may seem because the analytic continuation inherent in our earlier discussion of smooth boundaries generally leads to singularities in complex space. However we find some quite unexpected results concerning the Sommerfeld problem in which halflines are illuminated by real or complex plane waves or Gaussian beams.

Because of the difficulties associated with multiple reflections, we do not consider interior problems for the Helmholtz equation in this article. However, since the complexification say, of a one-dimensional surface in two-dimensional space is a two-dimensional surface in four-dimensional space, there is no distinction in principle between interior and exterior problems, as far as complex ray theory is concerned.

2 Ray Theory

The way in which classical ray theory can be used to find high-frequency approximate solutions to the Helmholtz equation has been described in many places [24, 1, 27] (and references given therein). The ray equations for scalar, time-harmonic wave propagation with real wavenumber k arise by applying a WKB ansatz

$$\phi \sim e^{iku(\boldsymbol{x})} \sum_{n=0}^{\infty} \frac{i^n A_n(\boldsymbol{x})}{k^n} \tag{1}$$

to the Helmholtz equation

$$\nabla^2 \phi + k^2 N(\boldsymbol{x})^2 \phi = 0, \tag{2}$$

in the limit $k \to \infty$, where N is the refractive index. This results in the eikonal equation

$$(\nabla u)^2 = N(\boldsymbol{x})^2 \tag{3}$$

and the recursive system of transport equations

$$2\nabla A_0 \cdot \nabla u + A_0 \nabla^2 u = 0, \tag{4}$$

$$2\nabla A_n \cdot \nabla u + A_n \nabla^2 u = \nabla^2 A_{n-1}, \quad n \geq 1. \tag{5}$$

The system is closed by imposing suitable boundary conditions, including appropriate behaviour at infinity if necessary. Such data may be given on a surface S parametrised by $\boldsymbol{s}$, $\boldsymbol{x} = \boldsymbol{x}_0(\boldsymbol{s})$, and a radiation or extinction condition at infinity will select the branch of u that is physically appropriate.

In this paper we will almost exclusively be concerned with the case of a uniform, two-dimensional medium for which $N(\boldsymbol{x}) \equiv 1$, for ease of exposition (though extensions of our results to three dimensions and for variable refractive indices are possible). In this case, (2) and (3) become

$$\nabla^2 \phi + k^2 \phi = 0, \tag{6}$$

$$(\nabla u)^2 = 1, \tag{7}$$

and, if $u = u_0(s)$ is given on S, then the characteristic equations are

$$\frac{\partial p}{\partial t} = 0, \qquad \frac{\partial q}{\partial t} = 0, \qquad \frac{\partial x}{\partial t} = p, \qquad \frac{\partial y}{\partial t} = q, \qquad \frac{\partial u}{\partial t} = 1,$$

where $(p, q) = \nabla u$ and t represents distance along a characteristic. These may be integrated to give $p = p_0(s)$, $q = q_0(s)$ and

$$x = p_0(s)t + x_0(s), \tag{8}$$

$$y = q_0(s)t + y_0(s), \tag{9}$$

$$u = t + u_0(s), \tag{10}$$

where p_0 and q_0 are determined from the conditions

$$p_0^2 + q_0^2 = 1, \tag{11}$$

$$\frac{dx_0}{ds} p_0 + \frac{dy_0}{ds} q_0 = \frac{du_0}{ds}, \tag{12}$$

with whichever choice of branch is needed ultimately to give a physically acceptable solution.

Assuming s, t, u_0, p_0 and q_0 are real, we see that the characteristics, which are the "rays" of ray theory, are straight lines in this case. Clearly points where the Jacobian

$$J = \frac{\partial(x, y)}{\partial(s, t)}$$

vanishes will be of interest, since at these points the map between $(s,\, t)$ and $(x,\, y)$ cannot be inverted and there is therefore a singularity in u. We find

$$J = \frac{p_0'}{q_0}(t+\beta(s)), \tag{13}$$

where

$$\beta(s) = \frac{q_0}{p_0'}\left(x_0' q_0 - y_0' p_0\right), \tag{14}$$

and the prime represents d/ds. Hence, unless $p_0' \equiv 0$, which can only occur if the solution represents a plane wave, such a singularity will always exist, although it may not be in physical space. The non-invertibility of the map at $t = -\beta(s)$ corresponds to a caustic of the wavefield.

To see the implication of this singularity for the amplitude A_0, let us turn to the transport equations (4), (5), which we may write as ordinary differential equations along the rays:

$$\begin{aligned} 2\frac{\partial A_0}{\partial t} + \nabla^2 u\, A_0 &= 0, & (15)\\ 2\frac{\partial A_n}{\partial t} + \nabla^2 u\, A_n &= \nabla^2 A_{n-1}, \quad n \geq 1. & (16) \end{aligned}$$

From (8), (9) we find

$$\begin{aligned} \frac{\partial s}{\partial x} &= \frac{q_0^2}{p_0'(t+\beta)}, & (17)\\ \frac{\partial s}{\partial y} &= \frac{p_0^2}{q_0'(t+\beta)}, & (18)\\ \frac{\partial t}{\partial x} &= p_0 + \frac{p_0(y_0' - q_0'\beta)}{q_0'(t+\beta)}, & (19)\\ \frac{\partial t}{\partial y} &= q_0 + \frac{q_0(x_0' - p_0'\beta)}{p_0'(t+\beta)}. & (20) \end{aligned}$$

Hence

$$\nabla^2 u = \frac{\partial p}{\partial x} + \frac{\partial q}{\partial y} = \frac{\partial p}{\partial s}\frac{\partial s}{\partial x} + \frac{\partial q}{\partial s}\frac{\partial s}{\partial y} = \frac{1}{t+\beta} = \frac{\partial}{\partial t}(\log J). \tag{21}$$

Letting $a_0(s)$ be the boundary value of A_0, equation (15) then gives

$$A_0(s,t) = \left(\frac{\beta(s)}{t+\beta(s)}\right)^{1/2} a_0(s), \tag{22}$$

or equivalently

$$A_0(s,t)^2 J(s,t) = a_0(s)^2 j(s), \tag{23}$$

where $j(s) = J(s,0)$ is the boundary value of J; thus the singularity in the phase u at $t+\beta(s) = 0$ leads to an unbounded amplitude, and hence to a breakdown of the ray ansatz (1).

From (22) we see that another kind of singularity occurs when $\beta = 0$, i.e. at points where the caustic meets the boundary $t = 0$ and the corresponding amplitude is undefined. From (14) we see that $\beta = 0$ corresponds to $(x_0', y_0') \cdot (q_0, -p_0) = 0$, which implies that the ray impinging the boundary at (x_0, y_0) is tangential.

The use of matched asymptotic expansions to solve inner problems near such singularities and then to match back into the outer ray solution forms the basis of the famous Geometrical Theory of Diffraction (GTD)

[21]. Inner diffraction problems are also necessary at singular points of the boundary data, for example at corners.

We conclude this section by writing down a representation of A_n which is especially convenient for the asymptotic analysis of §4. Using (17)-(21) we can obtain (with some calculation) the following expression for the Laplacian of A_n:

$$\begin{aligned}\nabla^2 A_n &= \frac{\alpha}{\tau^2}\frac{\partial}{\partial\sigma}\left(\alpha\frac{\partial A_n}{\partial\sigma}\right)+\frac{2\alpha^2(\beta'-u_0')}{\tau^2}\frac{\partial^2 A_n}{\partial\sigma\partial\tau}\\ &\quad -\frac{\alpha^2(\beta'-u_0')}{\tau^3}\frac{\partial A_n}{\partial\sigma}+\frac{1}{\tau}\frac{\partial A_n}{\partial\tau}+\frac{\partial^2 A_n}{\partial\tau^2}\\ &\quad +\frac{\alpha\left(\alpha(\beta'-u_0')\right)'}{\tau^2}\frac{\partial A_n}{\partial\tau}+\frac{\alpha^2(\beta'-u_0')^2}{\tau}\frac{\partial}{\partial\tau}\left(\frac{1}{\tau}\frac{\partial A_n}{\partial\tau}\right),\end{aligned} \tag{24}$$

where $\alpha = p_0/q_0'$, $\sigma = s$ and $\tau = t+\beta(s)$. Hence the equation for A_n becomes

$$\begin{aligned}2\frac{\partial A_n}{\partial\tau}+\frac{A_n}{\tau} &= \frac{\alpha}{\tau^2}\frac{\partial}{\partial\sigma}\left(\alpha\frac{\partial A_{n-1}}{\partial\sigma}\right)+\frac{2\alpha^2(\beta'-u_0')}{\tau^2}\frac{\partial^2 A_{n-1}}{\partial\sigma\partial\tau}\\ &\quad -\frac{\alpha^2(\beta'-u_0')}{\tau^3}\frac{\partial A_{n-1}}{\partial\sigma}+\frac{1}{\tau}\frac{\partial A_{n-1}}{\partial\tau}+\frac{\partial^2 A_{n-1}}{\partial\tau^2}\\ &\quad +\frac{\alpha\left(\alpha(\beta'-u_0')\right)'}{\tau^2}\frac{\partial A_{n-1}}{\partial\tau}+\frac{\alpha^2(\beta'-u_0')^2}{\tau}\frac{\partial}{\partial\tau}\left(\frac{1}{\tau}\frac{\partial A_{n-1}}{\partial\tau}\right).\end{aligned} \tag{25}$$

With the ansatz

$$A_n = \sum_{j=n}^{3n}\frac{g_{n,j}(s)}{\tau^{j+1/2}}, \tag{26}$$

where the upper and lower summation limits are calculated by considering the highest and lowers powers of $1/\tau$ occurring on each side of (25), we find that

$$\nabla^2 A_n = \sum_{j=n+2}^{3n+4}\frac{\gamma_{n,j}(s)}{\tau^{j+1/2}}, \tag{27}$$

where

$$\begin{aligned}\gamma_{n,j} &= \alpha(\alpha g_{n,j-2}')'+(j-3/2)^2 g_{n,j-2}\\ &\quad +2\alpha^2(\beta'-u_0')(j-3)g_{n,j-3}'-\alpha(\alpha(\beta'-u_0'))'(j-5/2)g_{n,j-3}\\ &\quad +\alpha^2(\beta'-u_0')^2(j-7/2)(j-3/2)g_{n,j-4}.\end{aligned} \tag{28}$$

Hence the transport equations (16) are satisfied providing $g_{n,j}$ satisfies the recurrence relation

$$\begin{aligned}g_{n+1,j} &= -\frac{1}{2j}\left[\alpha(\alpha g_{n,j-1}')'+(j-1/2)^2 g_{n,j-1}\right.\\ &\quad +2\alpha^2(\beta'-u_0')(j-2)g_{n,j-2}'-\alpha(\alpha(\beta'-u_0'))'(j-3/2)g_{n,j-2}\\ &\quad \left.+\alpha^2(\beta'-u_0')^2(j-5/2)(j-1/2)g_{n,j-3}\right].\end{aligned} \tag{29}$$

However, A_n as given by (26) is not the complete solution since it does not satisfy the boundary condition. To make it do so at each stage we must add on a solution to the homogeneous version of (16), a procedure which leads to

$$A_n = \frac{1}{\sqrt{\tau}}\sum_{m=0}^{n}\sum_{j=m}^{3m}\frac{g_{n,m,j}(s)}{\tau^j}, \tag{30}$$

where the new summation coefficients $g_{n,m,j}$ satisfy the recurrence relation

$$\begin{aligned} g_{n+1,m+1,j} &= -\frac{1}{2j}\Big[\alpha\left(\alpha g'_{n,m,j-1}\right)' + (j-1/2)^2 g_{n,m,j-1} \\ &\quad + 2\alpha^2(\beta'-u_0')(j-2)g'_{n,m,j-2} - \alpha\left(\alpha(\beta'-u_0')\right)'(j-3/2)g_{n,m,j-2} \\ &\quad +\alpha^2\left(\beta'-u_0'\right)^2(j-5/2)(j-1/2)g_{n,m,j-3}\Big], \end{aligned} \tag{31}$$

with the boundary conditions

$$g_{0,0,0} = \beta^{1/2}a_0(s), \tag{32}$$

$$g_{n,0,0} = -\sum_{m=1}^{n}\sum_{j=m}^{3m}\frac{g_{n,m,j}}{\beta^j} + \beta^{1/2}a_n(s). \tag{33}$$

This form of the solution for A_n is particularly suitable for the techniques we develop and employ in §4.

2.1 Ray theory versus Fourier Transform

Consider the following boundary value problem:

$$\nabla^2\phi + k^2\phi = 0 \qquad \text{in } y>0, \tag{34}$$

$$\phi = A_0(x)e^{iku_0(x)} \qquad \text{on } y=0, \tag{35}$$

together with some specification of the behaviour as $x^2+y^2\to\infty$, which will depend on the nature of u_0. We may approximate the solution with the ray equations (8)-(12) to give

$$x = s + u_0'(s)t, \tag{36}$$

$$y = t\sqrt{1-(u_0'(s))^2}, \tag{37}$$

$$u = u_0(s)+t. \tag{38}$$

Alternatively, we may solve exactly using the Fourier transform to give

$$\phi(x,y) = \frac{1}{2\pi}\int\left(\int A_0(s)e^{iku_0(s)+isv}\,ds\right)e^{i\sqrt{k^2-v^2}y-ixv}\,dv, \tag{39}$$

along some suitable contours, and for a specified branch of $\sqrt{k^2-v^2}$. Let us briefly review the relationship between these two approaches. Setting $v=kw$ in (39) gives

$$\phi(x,y) = \frac{k}{2\pi}\int\left(\int A_0(s)e^{ik(u_0(s)+sw)}\,ds\right)e^{ik(\sqrt{1-w^2}y-xw)}\,dw. \tag{40}$$

In many situations the dominant contribution to the first integral as $k\to\infty$ comes from the saddle point $s=s_0(w)$, where $w=-u_0'(s_0(w))$. Applying the method of steepest descents gives

$$\phi(x,y)\sim\int K(w)e^{ikf(w)}\,dw \tag{41}$$

where

$$f(w) = u_0(s_0(w)) + ws_0(w) + \sqrt{1-w^2}y - wx, \tag{42}$$

$$K(w) = \mu_1 A_0(s_0(w))\sqrt{\frac{k}{2\pi u_0''(s_0(w))}}, \tag{43}$$

and μ_1 is a constant independent of s_0 and k. The dominant contribution to ϕ as $k \to \infty$ therefore comes from the saddle point $w = w_0$, where

$$s_0(w_0) - x - \frac{w_0 y}{\sqrt{1 - w_0^2}} = 0, \tag{44}$$

or equivalently

$$s_0(w_0) - x + \frac{u_0'(s_0(w_0))y}{\sqrt{1 - (u_0'(s_0(w_0)))^2}} = 0. \tag{45}$$

This may be compared with the result of eliminating t from the ray solution (36), (37), viz.

$$s - x + \frac{u_0'(s)y}{\sqrt{1 - (u_0'(s))^2}} = 0. \tag{46}$$

Hence we see that the point of emergence from the boundary (s, say) of the ray passing through (x, y) is related to the saddle points of the double integral in (40) through $s = s_0(w_0)$. Finally, applying the method of steepest descents to (41) and using this result we find

$$\phi \sim \mu_2 K(w_0) \sqrt{\frac{2\pi}{k f''(w_0)}} e^{iku} \tag{47}$$

where

$$\begin{aligned} u &= u_0(s) - s u_0'(s) - \sqrt{1 - (u_0'(s))^2} y + u_0'(s) x, & (48) \\ &= u_0(s) + t, & (49) \end{aligned}$$

and μ_2 is another constant independent of k. Thus the ray approximation to the Helmholtz equation is often equivalent to an application of the method of steepest descents applied to the Fourier transform solution, should one exist. We now see one way of thinking of complex rays[1]. A complex value of s, corresponding to a ray coming from a point on the complexified boundary, corresponds to a complex saddle point in the first integral in (40).

2.2 Caustics

Since caustics are ubiquitous features of ray theory, and since they are the simplest example of the breakdown of a ray approach, we now examine their structure in more detail. This analysis will also provide a motivation for our systematic treatment of complex rays to follow in §3.

Consider the ray solution (considered previously in [20])

$$\begin{aligned} x &= \cos s - t \sin s, & (50) \\ y &= \sin s + t \cos s, & (51) \\ u &= s + t. & (52) \end{aligned}$$

As with all the examples to be considered subsequently in this paper, this solution implies that u is a multivalued function of (x, y) and, in fact, there are two real rays through any point in $x^2 + y^2 > 1$. The rayfield may be identified with the exact solution

$$\phi(r, \theta) = J_k(kr) e^{ik\theta} \tag{53}$$

[1] We will mention complex rays several times in the following pages but we defer a precise definition until §3.

of the Helmholtz equation, where k is a large integer and r, θ are plane polar coordinates. We may consider such a solution to arise from an incident field imposed at infinity or as corresponding to a trapped mode in a circularly cylindrical duct; the geometry is shown in Fig. 2.

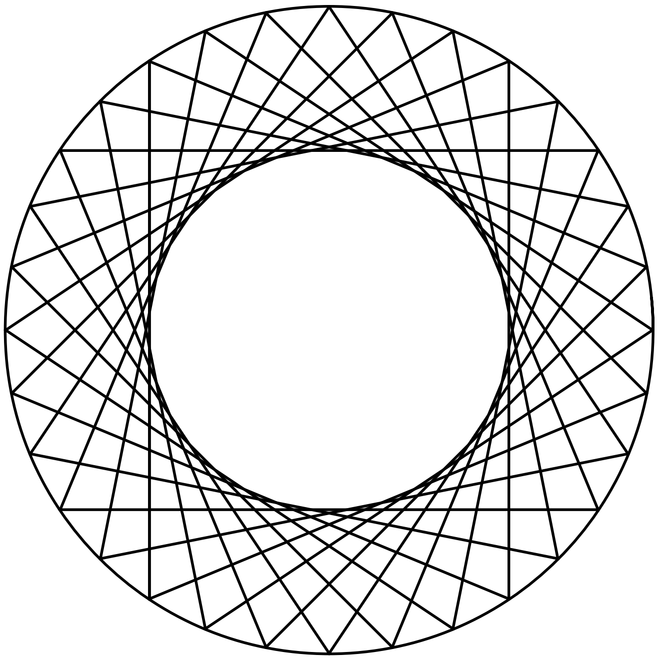


Figure 2: Trapped mode in a duct whose rays form a circular caustic.

It is not difficult to show that in this case $\alpha = 1$, $\beta = 0$, so that the curve $t = 0$, $x = \cos s$, $y = \sin s$ is a caustic of the ray field. The leading-order amplitude is given by $A_0 = 1/(-t)^{1/2}$, which as we have already seen, tends to infinity as the caustic is approached. Hence the ray expansion (50)-(52) is only valid for $t < 0$ (i.e. before a given ray reaches the caustic), and, to determine the "outgoing" wavefield that results on passing the caustic, it is necessary to consider a local problem in the vicinity of the caustic and match this to the incoming and outgoing wavefields, as described in [8, 31].

The symmetry of the problem means that we can invert the map (50), (51) between $(s,\ t)$ and $(x,\ y)$ to give

$$
\begin{aligned}
t &= -\sqrt{r^2-1}, & (54)\\
s &= \theta + \cos^{-1}(1/r), & (55)
\end{aligned}
$$

where r and θ are polar coordinates, and $r > 1$. Thus u and A_0 are given by

$$
\begin{aligned}
u(r,\theta) &= \theta + \cos^{-1}(1/r) - \sqrt{r^2-1}, & (56)\\
A_0(r,\theta) &= \frac{1}{(r^2-1)^{1/4}}. & (57)
\end{aligned}
$$

We introduce local coordinates near the caustic by writing

$$
r = 1 + \epsilon\rho, \tag{58}
$$

where the small parameter ϵ has yet to be determined. Inserting (58) into (56), (57) and expanding gives

$$u \sim \theta - \frac{2\sqrt{2}}{3}\epsilon^{3/2}\rho^{3/2} + \cdots, \tag{59}$$

$$A_0 \sim \frac{1}{\epsilon^{1/4}2^{1/4}\rho^{1/4}} + \cdots, \tag{60}$$

as $\epsilon \to 0$. Writing the Helmholtz equation in terms of these inner coordinates gives

$$\frac{1}{\epsilon^2}\frac{\partial^2 \phi}{\partial \rho^2} + \frac{1}{\epsilon(1+\epsilon\rho)}\frac{\partial \phi}{\partial \rho} + \frac{1}{(1+\epsilon\rho)^2}\frac{\partial^2 \phi}{\partial \theta^2} + k^2\phi = 0. \tag{61}$$

Motivated by the expansion (59) we seek an inner solution of the form $\phi = f(\rho)e^{ik\theta}$, giving

$$\frac{1}{\epsilon^2}\frac{\partial^2 f}{\partial \rho^2} + \frac{1}{\epsilon}(1 - \epsilon\rho + \cdots)\frac{\partial f}{\partial \rho} + k^2(2\epsilon\rho + \cdots)f = 0. \tag{62}$$

A sensible balance of terms arises when $\epsilon = k^{-2/3}$, giving Airy's equation

$$\frac{\partial^2 f}{\partial \rho^2} + 2\rho f = 0, \tag{63}$$

with general solution

$$f = C\,\mathrm{Ai}(-2^{1/3}\rho) + D\,\mathrm{Bi}(-2^{1/3}\rho). \tag{64}$$

The constants C and D are determined by matching with the outer region on the inside and outside of the caustic as follows:

(i) as $\rho \to -\infty$, $\mathrm{Ai}(-2^{1/3}\rho)$ is exponentially small, while $\mathrm{Bi}(-2^{1/3}\rho)$ is exponentially large. Since we require the solution to be small in the interior of the caustic, we therefore require that $D = 0$.

(ii) as $\rho \to \infty$,

$$\mathrm{Ai}(-2^{1/3}\rho) \sim -\frac{e^{-i\pi/4}}{2^{13/12}\sqrt{\pi}}\left(\frac{e^{-i2\sqrt{2}\rho^{3/2}/3}}{\rho^{1/4}} - \frac{e^{i\pi/2}e^{i2\sqrt{2}\rho^{3/2}/3}}{\rho^{1/4}}\right).$$

The first of these two terms matches into the incoming ray given by (59), (60) and this gives the constant C as

$$C = -2^{5/6}\sqrt{\pi}e^{i\pi/4}.$$

The second term is therefore the outgoing ray field. We see that this ray matches exactly into the outgoing ray of the outer ray expansion, with only one slight technicality. The phase change of $\pi/2$ comes about from t going from negative to positive in the amplitude $A_0 = 1/(-t)^{1/2}$. Matching with the inner region tells us which branch of the square root to take for $t > 0$.

Returning to the ray solution (50)-(51) we note that there are no real rays passing through any interior point $x^2 + y^2 < 1$. From (54)-(55) we see that values of $r < 1$ correspond to complex values of s and t. For each interior point there are two values of s and t. Hence there are two possible complex rays—one yielding an unphysical exponentially large amplitude which is discarded, and the other providing the exponentially small contribution which matches into the inner caustic region through (i) above.

2.3 Motivation for Complex Rays

The purpose of this section is to show that, as in the case of caustics, there are many practical problems in which the notion of a complex ray arises naturally. We shall do this using a number of examples from acoustic wave propagation.

Subsonic surface ray excitation A surface ray is one which is confined near the boundary of a wavebearing medium. An example is a creeping ray, initiated as a result of a ray at grazing incidence on a scatterer (to be discussed later), or a "free-mode", which is a solution of the homogeneous versions of both field equation and associated boundary condition. We discuss this second case by considering the homogeneous "impedance" problem

$$\begin{aligned}
(\nabla^2 + k^2)\phi &= 0, \qquad y > 0, && (65)\\
\frac{\partial \phi}{\partial y} + \mu k \phi &= 0, \qquad y = 0, && (66)\\
\phi &\to 0 \qquad \text{as } y \to \infty, && (67)
\end{aligned}$$

where $\mu > 0$ is an $O(1)$ constant. The plane wave solution of this system is

$$\phi = A \exp\left(ik\sqrt{1+\mu^2}x - \mu k y\right), \qquad y > 0, \tag{68}$$

for any constant A, which is a wave propagating without decay along the boundary $y = 0$ with a phase speed less than the wave-speed of the bulk medium and strong decay in $y > 0$. The corresponding ray solution is

$$\begin{aligned}
x &= \sqrt{1+\mu^2}t + s, && (69)\\
y &= i\mu t, && (70)\\
u &= \sqrt{1+\mu^2}s + t, && (71)
\end{aligned}$$

and the only rays to intersect the physical domain $(x, y) \in \mathbb{R}^2$, $y > 0$ necessarily emerge from complex points on the boundary (i.e. $s \in \mathbb{C}$) and propagate a complex distance before reaching the prescribed observation point (i.e. $t \in \mathbb{C}$). In fact, the ray passing through (x, y) emerges from the boundary point $s = x + iy\sqrt{1+\mu^2}/\mu$ and propagates a "distance" $t = -iy/\mu$. Hence we see straightaway that this ray structure can only be discerned by analytic continuation of s, t, x, y and the boundary data.

Analytic continuation is always a dangerous process and even in this trivial situation we can encounter singular behaviour when we complexify. Indeed, the plane-wave reflection coefficient R for a wave incident at an angle α to the normal is

$$R = \frac{i\cos\alpha - \mu}{i\cos\alpha + \mu}. \tag{72}$$

For all real angles of incidence R is finite and of unit modulus, implying that all the incident energy is radiated back into the bulk medium. Now suppose the boundary is insonified not by a plane wave, but by the field emitted from a line source located at the point (x_0, y_0). If this source is many wavelengths distant from the boundary, the problem can be analysed asymptotically in real space, using a locally planar approximation for the wavefronts near the boundary (see [42]). However, such an analysis will not take into account exponentially small terms which arise from the complex rays emitted from the source. These rays correspond to the outgoing source $H_0^{(1)}(kr)$ where $r = \sqrt{(x-x_0)^2 + (y-y_0)^2}$ is complex and they are given simply by $\theta = \tan^{-1}(y-y_0)/(x-x_0)$ =constant. However, when θ is such that $\cos\alpha = i\mu$, R is infinite and, according to [23], this critically incident complex ray is precisely the one which excites the surface ray described in the earlier part of this example. This surface field propagates to other parts of the boundary, including the physical domain, and so its initiation is of considerable practical importance and it demands the solution of a complexified diffraction problem as in [39]. Indeed, this study gives a crucial clue to the importance of what we will call Stokes switching, and we will return to it at the end of §4.4.

We remark that although our presentation of the example above has required the ill-posed procedure of analytic continuation, this does not contradict the well-posedness of the underlying boundary value problem for the Helmholtz equation. The situation is analogous to that encountered when contour integration is used to evaluate real definite integrals; many analytic continuations of the integrand may be possible but, no matter what their singularity structure, the consequences of Cauchy's theorem on the real axis will always be the same.

Total Internal Reflection Another familiar phenomenon which illustrates the need for a complex ray analysis is that of total internal reflection. Consider, for example, two semi-infinite, compressible fluids separated by an interface $y = 0$. Denoting the upper and lower fluids by the suffices 1 and 2 respectively, the problem of plane wave incidence is modelled by (Fig. 3)

$$(\nabla^2 + k_1^2)\phi_1 = 0, \qquad y > 0, \tag{73}$$

$$(\nabla^2 + k_2^2)\phi_2 = 0, \qquad y < 0, \tag{74}$$

$$\phi_1 = \phi_2, \qquad y = 0, \tag{75}$$

$$\frac{\partial \phi_1}{\partial y} = \lambda \frac{\partial \phi_2}{\partial y}, \qquad y = 0, \tag{76}$$

where k_i is the wavenumber in fluid i and λ is the ratio of the fluid density in region 1 to that in region 2.

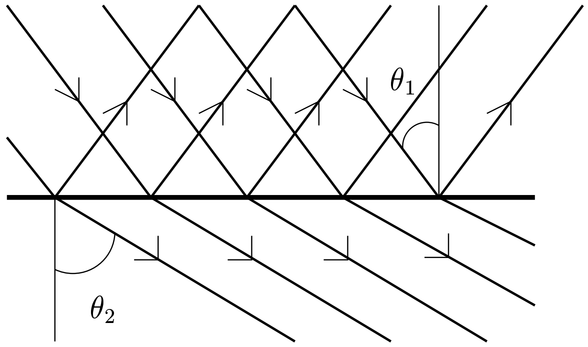


Figure 3: Reflection of a plane wave at an interface between two compressible media.

When we subtract out a prescribed incoming field

$$\phi_1^{(\mathrm{inc})}(x,y) = \exp[ik_1(x\sin\theta_1 - y\cos\theta_1)], \qquad 0 \le \theta_1 \le \pi/2, \tag{77}$$

and write $\phi_1 = \phi_1^{(\mathrm{inc})} + \phi_1^{(\mathrm{ref})}$, then we require that $\phi_1^{(\mathrm{ref})}$ is outgoing in $y > 0$ and a radiation or extinction condition must also be applied to ϕ_2 in $y < 0$. It is the latter condition that gives rise to complex rays in this problem, because, incorporating all the conditions at infinity, the solution turns out to be

$$\phi_1^{(\mathrm{ref})}(x,y) = \frac{n\cos\theta_1 - \lambda\cos\theta_2}{n\cos\theta_1 + \lambda\cos\theta_2}\exp[ik_1(x\sin\theta_1 + y\cos\theta_1)], \qquad y > 0, \tag{78}$$

$$\phi_2(x,y) = \frac{2n\cos\theta_1}{n\cos\theta_1 + \lambda\cos\theta_2}\exp[ik_2(x\sin\theta_2 - y\cos\theta_2)], \qquad y < 0, \tag{79}$$

where $n = k_1/k_2$ and $k_1 \sin\theta_1 = k_2 \sin\theta_2$, all assuming $k_2 > k_1$. However, if $k_2 < k_1$ there is the possibility that $\sin\theta_2 = (k_1/k_2)\sin\theta_1$ exceeds unity, giving rise to a *complex* angle of transmission θ_2. In this case we must choose the sign of the y-dependence in the exponent in (79) in order to guarantee decay as $y \to -\infty$. From a ray perspective, eikonal data for $\phi_1^{(\mathrm{ref})}$ and ϕ_2 are

$$p_1^{(\mathrm{ref})}(s,0) = \sin\theta_1, \qquad q_1^{(\mathrm{ref})}(s,0) = \cos\theta_1, \tag{80}$$

and

$$p_2(s,0) = \frac{k_1}{k_2}\sin\theta_1, \qquad q_2(s,0) = \left(1 - \frac{k_1^2}{k_2^2}\sin^2\theta_1\right)^{1/2}, \tag{81}$$

where the sign in $q_1^{(\mathrm{ref})}$ is chosen to satisfy the radiation condition in $y > 0$, but the branch for q_2 in (81) has yet to be specified. The ray equations in $y < 0$ are

$$x = \frac{k_1}{k_2}\sin\theta_1 t + s, \tag{82}$$

$$y = \left(1 - \frac{k_1^2}{k_2^2}\sin^2\theta_1\right)^{1/2} t, \tag{83}$$

$$u_2 = \frac{k_1}{k_2}\sin\theta_1 s + t, \tag{84}$$

from which we can see that if $k_2 < k_1$, then we must define

$$\left(1 - \frac{k_1^2}{k_2^2}\sin^2\theta_1\right)^{1/2} = -i\left(\frac{k_1^2}{k_2^2}\sin^2\theta_1 - 1\right)^{1/2}$$

in order to get decay as $y \to -\infty$. Thus, the expression

$$s = x - \frac{ik_1 \sin\theta_1\, y}{\sqrt{k_1^2\sin^2\theta_1 - k_2^2}} \tag{85}$$

links any observation point P in $y < 0$ to the complex point of emergence s on the boundary of the transmitted ray through P (as in Fig. 4).

What is happening is that real initiation points on the boundary ($s \in \mathbb{R}$) produce transmitted rays which immediately leave the physical domain. It is only those rays transmitted from that portion of the complexified boundary defined by (85) for $(x, y) \in \mathbb{R}^2$, $y < 0$, which provide the refracted field.

In the harder example where the configuration is forced by a remote, time-harmonic line source in the upper medium, the incoming field can be decomposed into a family of rays all emanating from the source point. As these rays impinge upon the boundary, they behave on a local scale like an incoming plane wave and all the previous arguments apply. That is, if $k_1 > k_2$ then those incident rays for which the angle of incidence θ satisfies $(k_1/k_2)\sin\theta < 1$ will reflect and transmit in the usual fashion. However, for those such that $(k_1/k_2)\sin\theta > 1$, the locally refracted rays instantly leave the physical domain, never to reappear. The associated transmitted field is then generated by a complex ray refracted from other points on the complexified boundary, as determined by the prescription outlined earlier. Some very interesting features of the evanescent transmitted field that occurs in this case will be discussed in §6.2. Apart from illustrating the fact that complex ray contributions can and do arise in apparently "real ray" scenarios, this second example also shows the need to examine *all* rays, real and complex, emitted from any source, since rays of the latter type might induce further complex rays at points of reflection or diffraction which then go on to have physical significance.

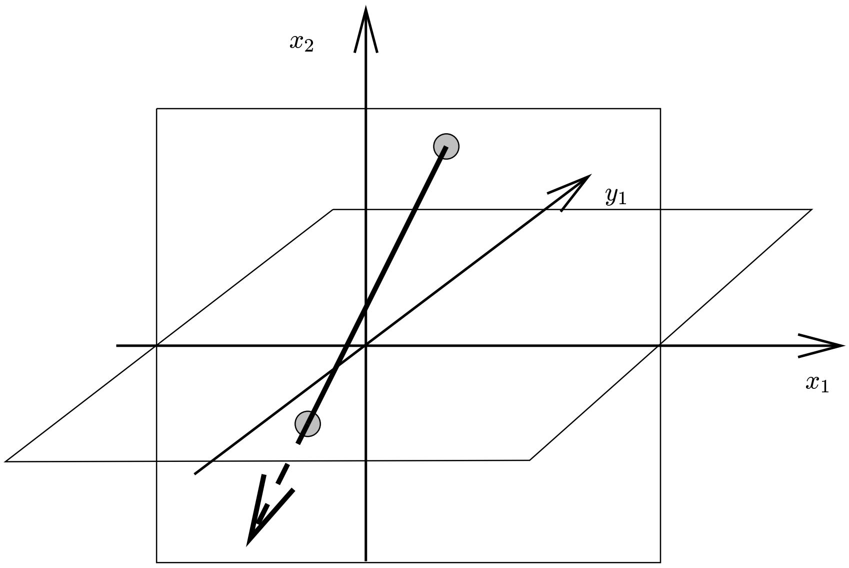


Figure 4: Three-dimensional cross-section of complex (x, y) space defined by $\mathrm{Im}\, y = 0$. Here $x = x_1 + ix_2$, $y = y_1$, where x_1, x_2, $y_1 \in \mathbb{R}$. The diagram shows a ray leaving the complexified boundary (the $x_1 - x_2$ plane) and intersecting real space (the (x_1, y_1) plane) in a point.

Gaussian Beams We define a Gaussian beam to be a collimated wavefield propagating in a specified direction with an amplitude which decreases *algebraically* along the direction of propagation and *exponentially* in a transverse direction with a decay exponent that is *quadratic* in the cross-profile coordinate. There is no such explicit exact solution to the Helmholtz equation[2] but it can be associated with an asymptotic representation of an exact solution solution valid in the so-called paraxial region close to the central axis of the beam [11]. The basic notion is to take the Green's function for the two-dimensional Helmholtz equation, $H_0^{(1)}(kr)$ where $r = \sqrt{x^2 + y^2}$, and to shift the origin to a suitably defined complex location.

Suppose the source is located at the point (x_0, y_0) with

$$x_0 = (b_1 - ib_2) \sin \alpha, \qquad y_0 = -(b_1 - ib_2) \cos \alpha \tag{86}$$

where b_1, $b_2 > 0$ and α are given real parameters. Identifying

$$\begin{aligned} x_1 &= x \cos \alpha + y \sin \alpha, & (87) \\ y_1 &= -x \sin \alpha + y \cos \alpha + b_1 & (88) \end{aligned}$$

allows the complex radial distance $r = \sqrt{(x - x_0)^2 + (y - y_0)^2}$ to be written in the form

$$r = \sqrt{x_1^2 + (y_1 - ib_2)^2}. \tag{89}$$

The paraxial region is now defined by

$$x_1^2 \ll y_1^2 + b_2^2 \tag{90}$$

[2]Indeed, there is no systematic theory concerning the circumstances in which a source (say a wave emitted from the end of a waveguide) can set up a Gaussian beam.

and if this condition is satisfied, then

$$r \sim y_1 - ib_2 + \frac{x_1^2}{2(y_1 - ib_2)}. \tag{91}$$

If, in addition to (90), we impose the "far-field" condition $ky_1 \gg 1$, then the modulus of the complex-valued argument of the Hankel function $H_0^{(1)}(kr)$ is uniformly large, and the Hankel function may be replaced by the first term in its asymptotic expansion, i.e., by $ie^{ikr}/(4\sqrt{kr})$. Incorporating (91) into this expression then yields an asymptotic field proportional to

$$\frac{e^{kb_2}}{\sqrt{y_1 - ib_2}\sqrt{k}} \exp\left[iky_1 + \frac{ikx_1^2}{2(y_1 - ib_2)}\right]$$

which has precisely the qualities required of a Gaussian beam. Thus, the y_1-axis is the principal axis of propagation and x_1 is the cross-profile coordinate. Geometrically, α is the angle of incidence of the beam, measured from the normal, with respect to the boundary $y = 0$, b_1 is the distance of the "waist" of the beam from the origin $x = 0$, $y = 0$, and $\sqrt{b_2/k}$ is the thickness of the beam at its waist (see Fig. 5). The lines in Fig. 5 represent the level curves of $\operatorname{Im} u$, and thus indicate lines of constant magnitude of the wavefield. These are also the normals to the level curves of $\operatorname{Re} u$, and can be shown to be the direction of energy flow to leading order.

Thus we see that as far as its specification in the physical domain is concerned, the beam has essentially four parameters: (a) angle of incidence, (b) width of beam, (c) distance of waist from the origin, and (d) specification that the beam passes through the origin. An alternative specification would be to replace the two conditions (c) and (d) by the coordinates of the "centre" of the beam, this being defined as the midpoint of the waistline. These four degrees of freedom are just the four degrees of freedom available to us in specifying the complex source coordinates x_0 and y_0.

Whilst we shall discuss in detail some problems involving the diffraction of Gaussian beams under a variety of circumstances in later parts of this paper, it is important to understand the ray structure of a Gaussian beam according to this prescription. The beam is generated by a uniform source emitting an isotropic family of complex rays from the complex source point defined by (86). The parameter that varies from one ray to the next is the polar angle, θ, measured with respect to this source point as origin. Thus, the equation of the rays can be written as $\theta =$ constant, where

$$\frac{y - y_0}{x - x_0} = \tan\theta.$$

If we now write $\tan\theta = t_1 + it_2$, t_1, $t_2 \in \mathbb{R}$, and substitute for x_0 and y_0 from (86), and restrict our attention to real values of x and y, then separating real and imaginary parts yields

$$y + b_1\cos\alpha = t_1(x - b_1\sin\alpha) - b_2 t_2 \sin\alpha, \tag{92}$$

$$-b_2\cos\alpha = t_1 b_2 \sin\alpha + t_2(x - b_1\sin\alpha). \tag{93}$$

These equations therefore allow t_1 and t_2 (and hence θ) to be calculated for a specified real observation point, showing that such a value of θ is generally complex. Generally, there is a one-to-one correspondence between (x, y) and (t_1, t_2), showing that the complex rays emitted from the source intersect the physical domain in *points*, at which the eikonal and transport equations can be solved to give the local field. The superposition of all these point contributions then forms the Gaussian profile. The exception that arises is

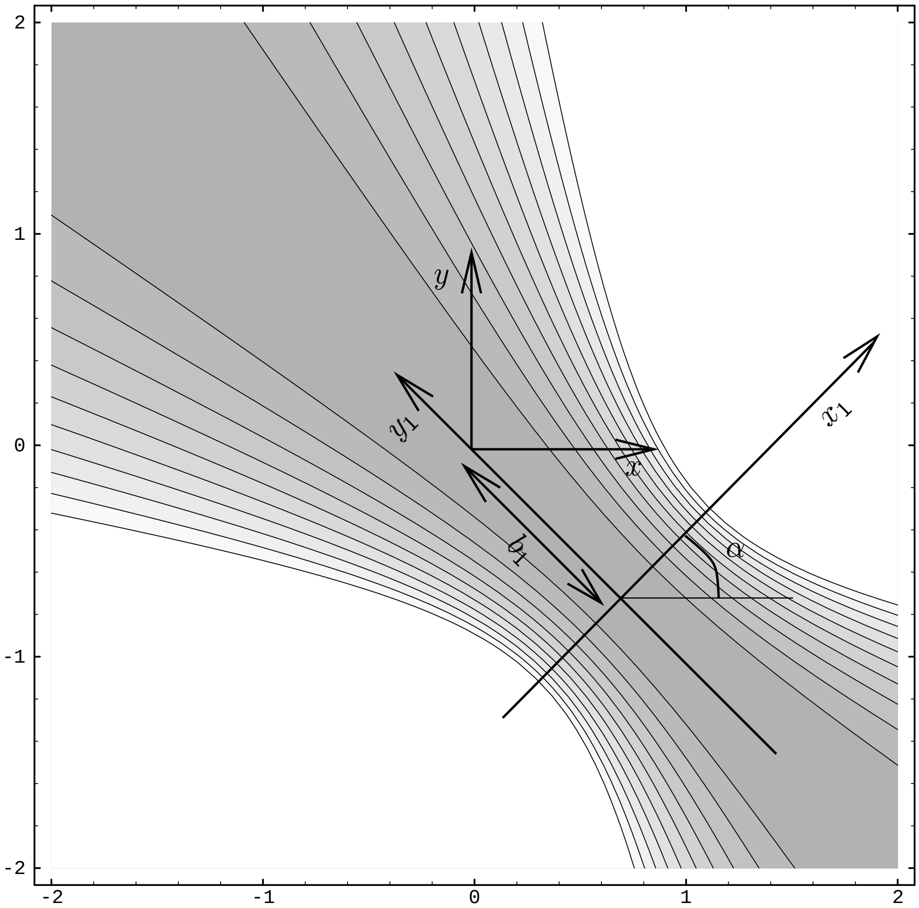


Figure 5: Diagram of a Gaussian Beam.

when $t_2 = 0$, so that θ is real (but note that θ is still measured with respect to a complex origin). In this case, $t_1 = -\cot\alpha$ and $y = -x\cot\alpha$. Thus it is for this value of θ, and only for this value of θ, that the associated complex ray intersects the physical domain along a straight line, and this line is coincident with the central axis of the beam.

The discussion above indicates the need for a careful geometric characterisation of complex rays, which we give in the next section.

3 The Geometry of Complex Rays

We found in the previous section that even in very simple two-dimensional ray problems, there may be real values of x and y which correspond to complex values of s and t. The idea behind complex ray theory is to consider all complex values of s, t, x and y from the beginning, and then, by considering all ray intersections with real space, enumerate all the relevant complex ray contributions to the solution of the physical problem. We choose the following definition:

Definition A *complex ray* is the set of complex points (x, y) corresponding to fixing a complex value of s in the ray equations (given by (8)–(12) in the two-dimensional case, with x_0, y_0 and u_0 necessarily analytic functions[3]) and then allowing the distance parameter t to range over all complex values, with $t = 0$ coinciding with the initiation point on the boundary.

Whatever the dimensionality of the ray problem in physical space, the equations of the rays can be written in the vector form

$$\boldsymbol{x} = \boldsymbol{p}_0(s)t + \boldsymbol{x}_0(s), \tag{94}$$

where $\boldsymbol{p}_0(s) = \nabla u$ and $\boldsymbol{x}_0(s)$ is a parametrisation of the boundary. Thus, fixing $s \in \mathbb{C}$ prescribes the initiation point $x_0(s)$ of the ray and its "direction of propagation" $\boldsymbol{p}_0(s)$, both of these quantities generally being complex. The dimensionality of the ray is therefore the number of degrees of freedom in t. Hence, a complex ray is always a two-dimensional manifold which exists in either the four-dimensional space of complex x and y or the six-dimensional space of complex x, y and z, depending upon the nature of the underlying problem. In the case we consider here of constant refractive index, this manifold is flat.

Whether or not a given complex ray intersects the physical domain can now be expressed geometrically in terms of the intersection (or not, as the case may be) of a two-dimensional plane (the complex ray) with either a two-dimensional plane (i.e. $\mathbb{R}^2$) if the physical problem is two-dimensional or with a three-dimensional plane if the physical problem is three-dimensional. This is likely to happen in the former case, but less likely in the latter, since there is an extra dimension in which the complex ray can be lost. In either case, if a complex ray does penetrate real space the generic intersection is in a point. The real rays of geometrical optics and GTD correspond to the degenerate case in which this intersection is a line.

One final general comment is that our definition implies that for given ray data away from points of diffraction, the ray leaving a point on the boundary is unique. For a given point on the ray, there is an infinity of one-dimensional curves lying within the ray linking that point to the initiation site on the boundary. Each such curve satisfies the ray equations in its own right, because it is embedded within a

[3]We will consider the case of nonanalytic x_0 and y_0 in §7.

higher-dimensional surface (which is the union of these curves) which does. We choose the latter for our definition of a complex ray, in agreement with [14, 13], whereas some authors (as in [26]) take the former.

As a typical example of complex ray tracing, we can return to the problem of the circular caustic, and ask where the rays intersecting $x^2 + y^2 < 1$ meet the caustic. These rays will be tangent at complex points on the caustic as we can see by considering the three-dimensional cross-section $\mathrm{Im}(y) = 0$ of four-dimensional space. Then the caustic $x^2 + y^2 = 1$ consists of the circle $x_2 = 0$, $x_1^2 + y_1^2 = 1$ and the hyperbola $x_1 = 0$, $y_1^2 - x_2^2 = 1$, where $x = x_1 + ix_2$, $y = y_1 + iy_2$, with x_1, x_2, y_1, y_2 real (see Fig. 6). Rays tangent to the circle are the real rays of §2.2. Rays tangent to the hyperbola intersect real space in a point within $x^2 + y^2 < 1$.

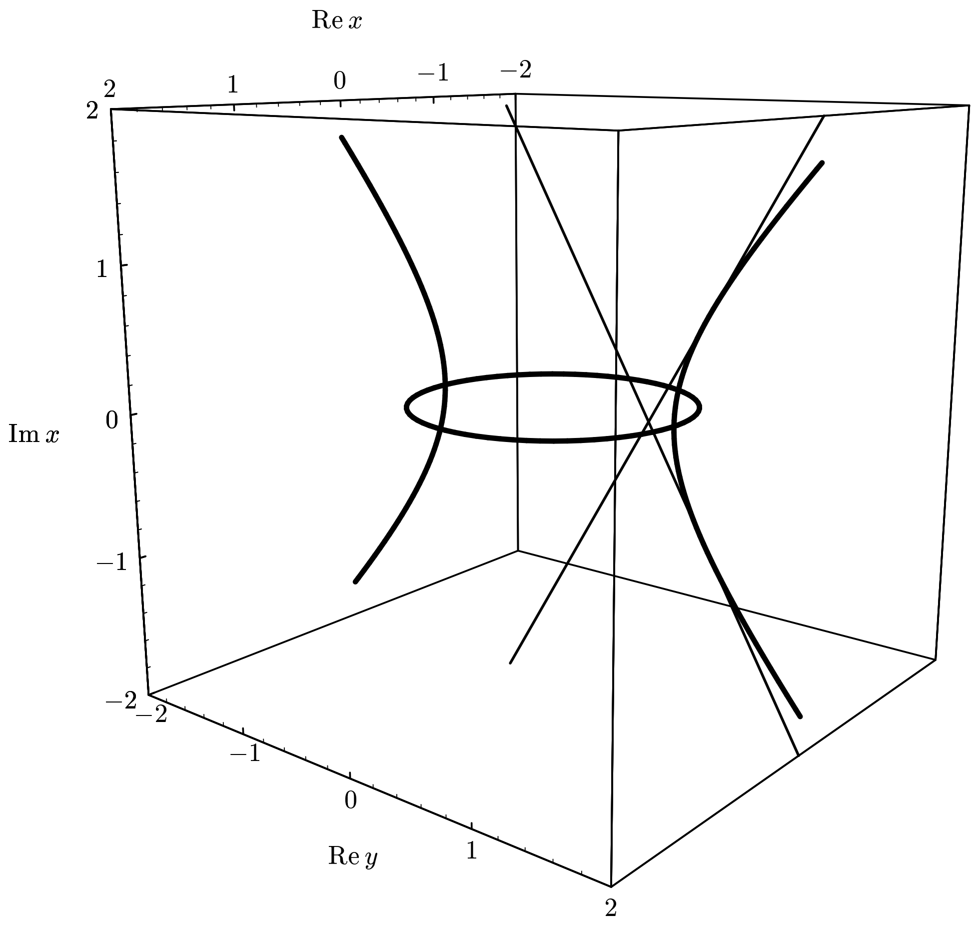


Figure 6: Three-dimensional cross-section of (x, y) space defined by $\mathrm{Im}\, y = 0$. The diagram shows rays tangent to complex points on the circular caustic $x^2 + y^2 = 1$ intersecting real space in the region $(\mathrm{Re}\, x)^2 + (\mathrm{Re}\, y)^2 < 1$.

Let us now make some further dimensional observations, based on two-dimensional physical problems, and raise some questions.

(i). The intersection between a propagating complex ray and a complexified boundary will also be a point in general. Thus the intersection of an incident complex ray with the boundary does not split the ray into two halves, as is the case with real rays, where the boundary divides the incident real ray into incoming

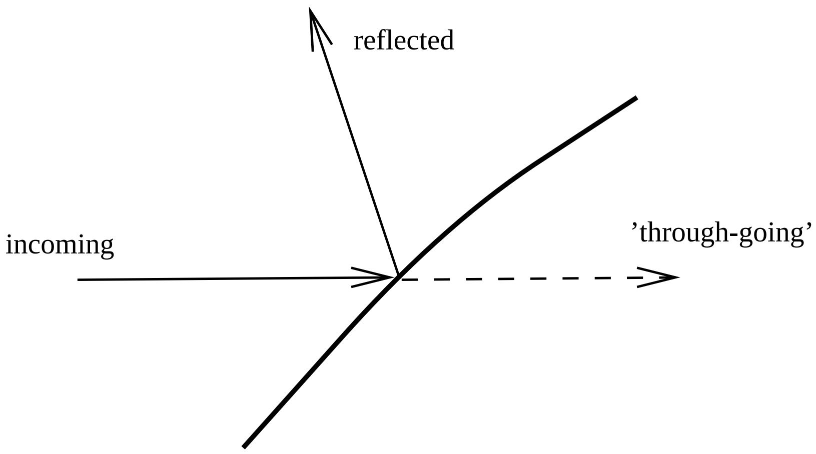


Figure 7: Diagram of the boundary of a scatterer dividing a ray into reflected and "through-going" components.

and "through-going" parts (see Fig. 7). The presence of the boundary means that the through-going ray is discarded. Such rules are not available for complex rays. Put another way, since the complexified boundary itself is a two-dimensional manifold in four-dimensional space, it has no inside or outside. In real ray problems the boundary is one-dimensional in two-dimensional space, and we solve on the outside of the boundary only. Indeed, analytically continuing this solution to the inside will generically lead to singularities in the solution since it involves solving the Cauchy problem inside a closed curve. In complex space we cannot separate the inside from the outside.

(ii). Following on from (i), a real ray at two-point grazing incidence to a smooth boundary will divide the boundary into "lit" and "unlit" regions, as shown in Fig. 8. However, because the intersection of a complex grazing ray with the complexified boundary is a point, and the complexified boundary is a two-dimensional manifold, the intersection cannot divide the complexified boundary into "lit" and "unlit" regions. How then are we to determine the lit portion of the complexified boundary?

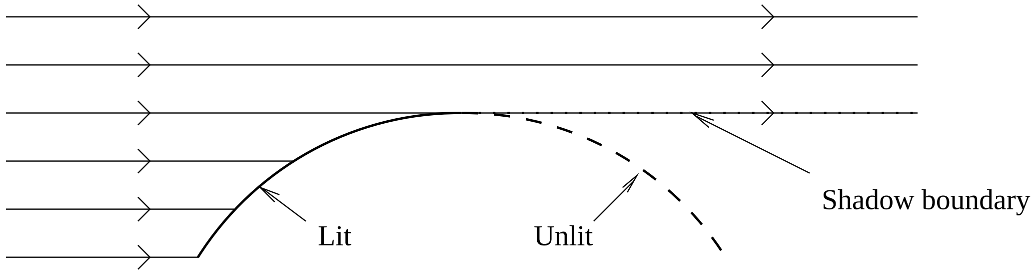


Figure 8: Diagram of a tangency point determining the lit portion of a scatterer and the tangent ray forming the shadow boundary

Moreover, for real rays, a ray at two-point grazing incidence forms the boundary of the shadow region; on one side of it the incident and reflected rays are present, on the other side of it they are not (see Fig. 8). In complex space, however, the tangent ray is only two-dimensional, and therefore cannot bound a region of four-dimensional space, so that it cannot be the boundary of the shadow region in complex space.

(iii). A complex caustic of the wavefield is a two-dimensional manifold, and therefore every ray meets the caustic in a point. Hence the caustic, like the complexified boundary, does not divide the ray into incoming and outgoing halves.

(iv). Finally, since the complex caustic is a two-dimensional manifold in four-dimensional space, it too cannot have an inside and outside as a real caustic does. What, then, do we mean by the dark side of a caustic?

Thus we see that once complex rays are being considered much of the intuition we have built up using real rays is lost. However, the observations and questions above also illustrate some of the advantages of complex rays. Let us first consider the implications of (iv).

Consider again the circular caustic at $x^2 + y^2 = 1$. Let us work for the moment in the $y = 0$ plane, so that the caustic lies at $x = \pm 1$. Then, if we are restricted to real space, we cannot get from the outside of the caustic ($|x| > 1$) to the inside of the caustic ($|x| < 1$) without passing through the caustic, and it is this which necessitated the matched asymptotic analysis of §2.2. However, if complex values of x are allowed, we can analytically continue the solution from the bright side around to the dark side without going near the singularities at $x = \pm 1$. Let us start with the *incoming* ray field

$$\begin{aligned} u &= \theta + \cos^{-1}(1/r) - \sqrt{r^2 - 1} \\ &= \cos^{-1}(1/x) - \sqrt{x^2 - 1}, \qquad (95) \\ A_0 &= \frac{1}{(r^2 - 1)^{1/4}} = \frac{1}{(x^2 - 1)^{1/4}}, \qquad (96) \end{aligned}$$

at A in Fig. 9, and analytically continue anti-clockwise around the caustic at $x = 1$. Then, on reaching real space on the inside of the caustic at B, u is positive imaginary and ϕ is exponentially small. Allowing complex rays means that we can compute the solution in the interior of the caustic. Note that although the analysis of §2.2 gave us the solution in the vicinity of the caustic, it did not give us an approximation to ϕ away from the caustic on the dark side. Here our analytic continuation gives us a recipe for this exponentially small field.

Proceeding with the analytic continuation we arrive back on the bright side at C. Here we are on the other branch of the square root in u, so we arrive at the *outgoing* ray $u = -\cos^{-1}(1/x) + \sqrt{x^2 - 1}$. The amplitude A_0 has changed by a factor $e^{-i\pi/2}$. Hence we have retrieved the outgoing wavefield complete with the correct phase change, and we have not had to consider the inner expansion near the singularity at $x = 1$ at all. This is an example of the "Keller-Maslov" theory of semiclassical mechanics [22]. In that theory it is also crucial to consider the phase function u as living on a Riemann surface in which the caustics are the branch points.

Our analytic continuation procedure has expected inherent dangers as we can see if we proceed further with the analytic continuation to arrive again on the inside of the caustic, because this time the branch of the square root will correspond to the exponentially large solution (D). Clearly we need some rule to tell us when to stop analytically continuing. Not all complex rays should be included in the solution at all points in space if we are to identify physically acceptable solutions.

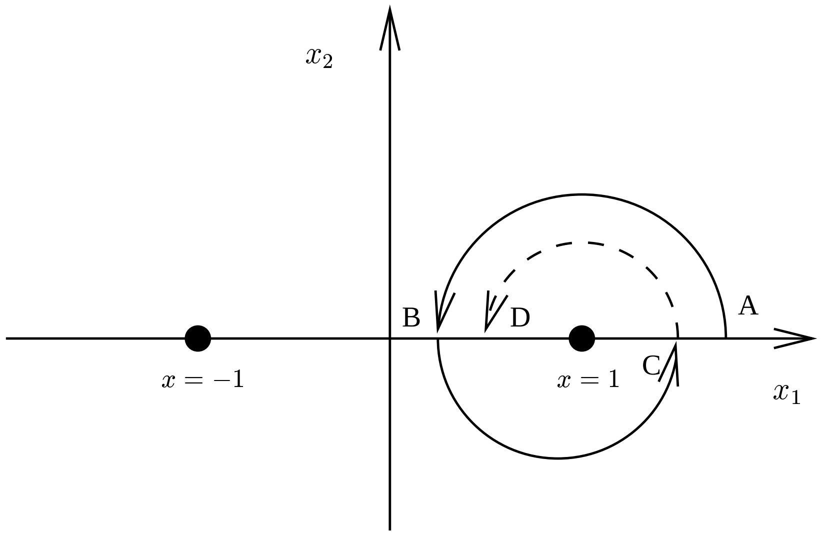


Figure 9: Diagram indicating the analytic continuation of the incoming ray solution around the caustic $x^2 + y^2 = 1$ in the plane $x = x_1 + ix_2$, $y = 0$, where x_1, $x_2 \in \mathbb{R}$.

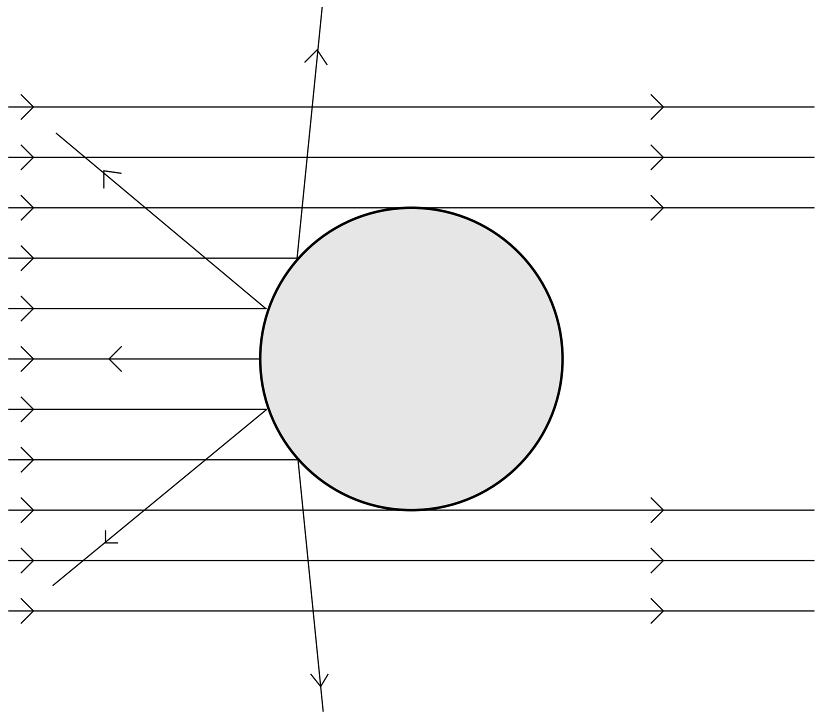


Figure 10: Incident and reflected rays when a plane wave is scattered by a circular cylinder.

This fact becomes even more obvious when we consider simple reflection problems, for example the reflection of a plane wave by a cylinder (see Fig. 10). If we analytically continue the incident and reflected wavefields and consider all complex values of s and t we immediately end up with the wavefields shown in Fig. 11, in which the incident field proceeds all the way through the cylinder to the shadow region and we have "reflected" rays coming in from infinity. These rays had previously been ruled out by the selection procedures for real rays and the laws of causality. However, as observed in (i)–(iv), these procedures cannot be applied to complex rays.

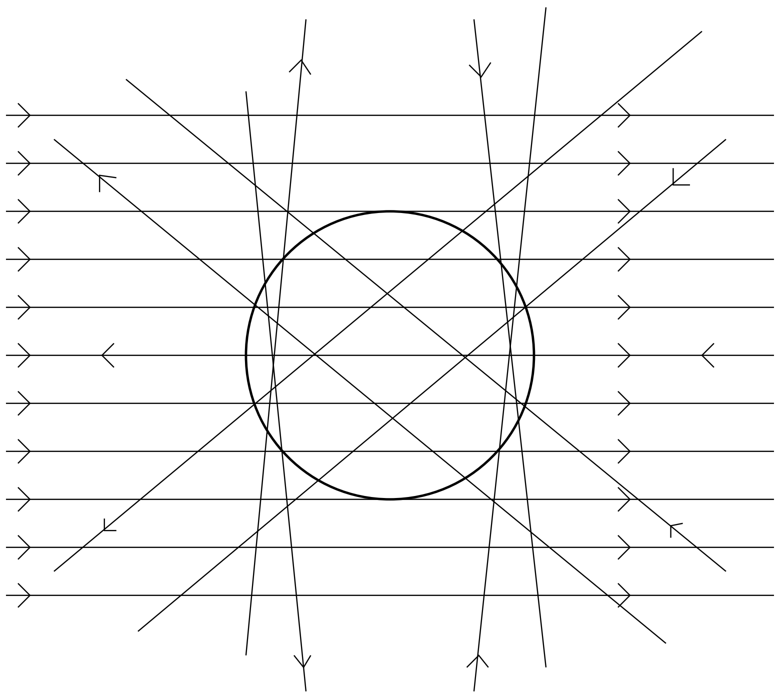


Figure 11: Rays obtained on analytically continuing the rays in Fig. 10.

In fact, the whole idea of causality itself is lost when considering complex rays. Since real rays correspond to $t \in \mathbb{R}$ there is a natural ordering of the points on the ray, which gives the ray a direction; in this way we can distinguish incoming and outgoing rays. Although it is possible to regard a complex ray as the union of one-dimensional manifolds on which $\mathrm{Re}\, t$ provides an ordering parameter [43], it is not clear that such a definition would necessarily lead to unique physically relevant solutions of the Helmholtz equation. Thus, since we can no longer distinguish incoming and outgoing rays and appeal to causality, we need a new set of rules for determining which complex rays are "active" (i.e. contribute to a physically meaningful rayfield) at any point in space.

4 Stokes' Phenomenon and Complex Rays

All our discussion thus far has led up to the basic question of the relationship between the process of analytic continuation and WKB asymptotic expansions. Hence it has provided the motivation for asking the question of how Stokes' phenomenon is involved in complex ray theory. To answer this question, three attributes of

Stokes' phenomenon must be borne in mind, and we begin by listing them.

4.1 Stokes' Phenomenon

Let us first briefly review the role of Stokes' phenomenon in the WKB approximation of ordinary differential equations [33, 36]. Consider, for example, the ordinary differential equation

$$\frac{d^2w}{dz^2} + 2z\frac{dw}{dz} = 0. \tag{97}$$

Applying a WKB ansatz for large $|z|$ leads to the "ray" solutions $w(z) \sim$ constant and $f(z) \sim e^{-z^2}/z$. Suppose we take the ray solution $w(z) \sim e^{-z^2}/(\sqrt{\pi}z)$ for large positive z and attempt to analytically continue by sweeping z around anti-clockwise in a large circular arc in the complex plane, as in Fig. 9. Then there is nothing to indicate the breakdown of the WKB ansatz, and we would arrive on the negative z axis with the same ray. However, we know that the solution to (97) with the behaviour $w(z) \sim e^{-z^2}/(\sqrt{\pi}z)$ as $z \to \infty$ is $w(z) = \operatorname{erfc}(z)$, and that $\operatorname{erfc}(z) \sim 2 + e^{-z^2}/(\sqrt{\pi}z)$ as $z \to -\infty$. Thus there is a multiple of the second ray present. This ray is "switched on" at the Stokes line $\arg z = \pi/2$, and is switched off again at the Stokes line $\arg z = 3\pi/2$. Such switching appears at first sight to be a jump in the amplitude of the ray in the WKB approximation, but is in fact a smooth but rapid transition as was found in [4], made rigorous in [32], and formulated in the language of matched asymptotic expansions in [34]. The Stokes lines are characterised as lines along which $w_1(z) = e^{-z^2}/(\sqrt{\pi}z)$ is *maximally* dominant over $w_2(z) = 2$ (this being the specific value of the constant here), or equivalently,

$$\text{the phases of } w_1 \text{ and } w_2 \text{ are equal,} \tag{98}$$

that is, $\operatorname{Im}(-z^2) = \operatorname{Im}(0)$. Thus the "ray" picture of the solution $\operatorname{erfc}(z)$ is as shown in Fig. 12. The ray $e^{-z^2}/(\sqrt{\pi}z)$ is always present, but the ray 2 only contributes in the left half-plane. In Fig. 12 we have also indicated the anti-Stokes lines, which we define as the lines on which the amplitudes of w_1 and w_2 are equal, that is, $\operatorname{Re}(-z^2) = \operatorname{Re}(0)$.

4.2 Stokes' phenomenon as saddles cross steepest descent paths

Consider next the integral representation of the complementary error function

$$w(z) = \operatorname{erfc}(z) = \frac{2}{\sqrt{\pi}}\int_z^\infty e^{-v^2}\,dv. \tag{99}$$

Applying the method of steepest descents to this integral gives the WKB approximation $w(z) \sim e^{-z^2}/(\sqrt{\pi}z)$ as $z \to \infty$. However, for this approximation to be valid, the contour of integration must now follow the steepest descent path from z to infinity. For $\operatorname{Re} z > 0$ this path is as shown in Fig. 13(a), but, when $\operatorname{Re} z < 0$ the steepest descent path from z goes to $-\infty$, and we need to add an extra integration along the real axis (see Fig. 13(b)). The extra ray contribution 2 arises from this integration.

This change in geometry of the steepest-descent path is typical of Stokes' phenomenon. Here it occurs because the steepest-descent path from z goes through the saddle at $v = 0$ when $\operatorname{Re} z = 0$.

Since the contribution from a saddle point is "switched on" when the steepest-descent path crosses it, we see why the Stokes lines should correspond to lines of equal phase of two rays, since at the point where

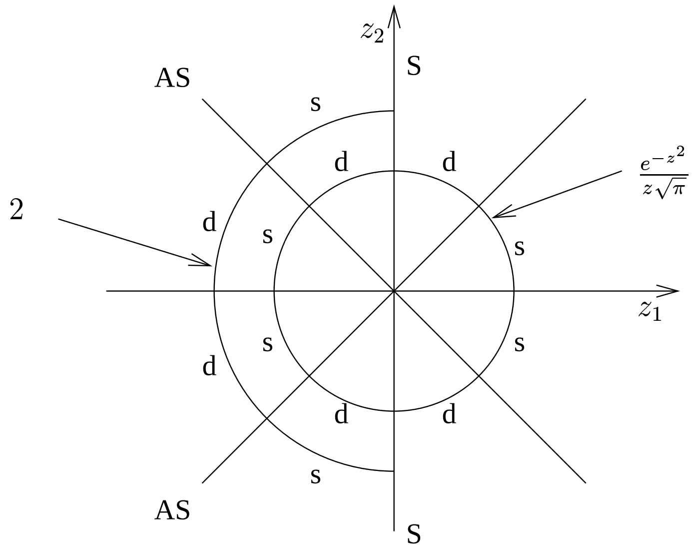
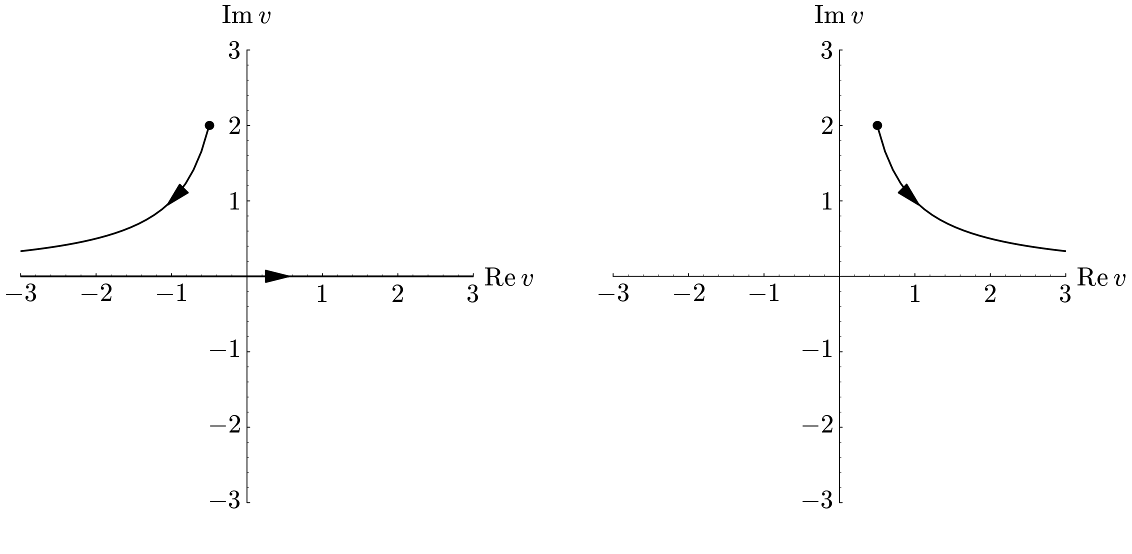


Figure 12: Diagram indicating the presence of the "rays" in the WKB expansion of the complementary error function; d and s indicate the dominant and subdominant rays respectively, and S and AS indicate the Stokes lines and anti-Stokes lines.



(a) $\mathrm{Re}\, z < 0$

(b) $\mathrm{Re}\, z > 0$

Figure 13: Steepest descent paths for $\mathrm{erfc}(z)$.

the extra saddle is picked up both contributing points (in this case, saddle and end-point) lie on the same steepest-descent (and therefore constant-phase) path. Of course, not all equal-phase lines will be Stokes lines, as is the case for $\arg z = 0$ and π in Fig. 12. Everything will depend upon the geometry of the saddle points and steepest descent contours. In general, the equal-phase lines will be easy to calculate, but the ability to discern the saddle-point geometry depends on the availability of a sufficiently simple integral representation of the solution. Thus we will see that one of our main difficulties will be deciding which of the equal-phase lines are Stokes lines.

Exactly the same scenario emerges for contour integrals involving singular integrands or with finite limits of integration. For example the integral

$$\int_0^{\pi} \frac{e^{izt}\,dt}{1+t^2} \tag{100}$$

studied in [35] has possible asymptotic contributions from $t = 0, \pi, \pm i$ as $|z| \to \infty$. Here the contributions from the poles $t = \pm i$ are switched on when the real parts of the phases of the end-point contributions $e^0, e^{i\pi z}$ are equal to those of the residues at the poles, $\mp(i/2)e^{\mp z}$. Hence the z-plane is as in Fig. 14.

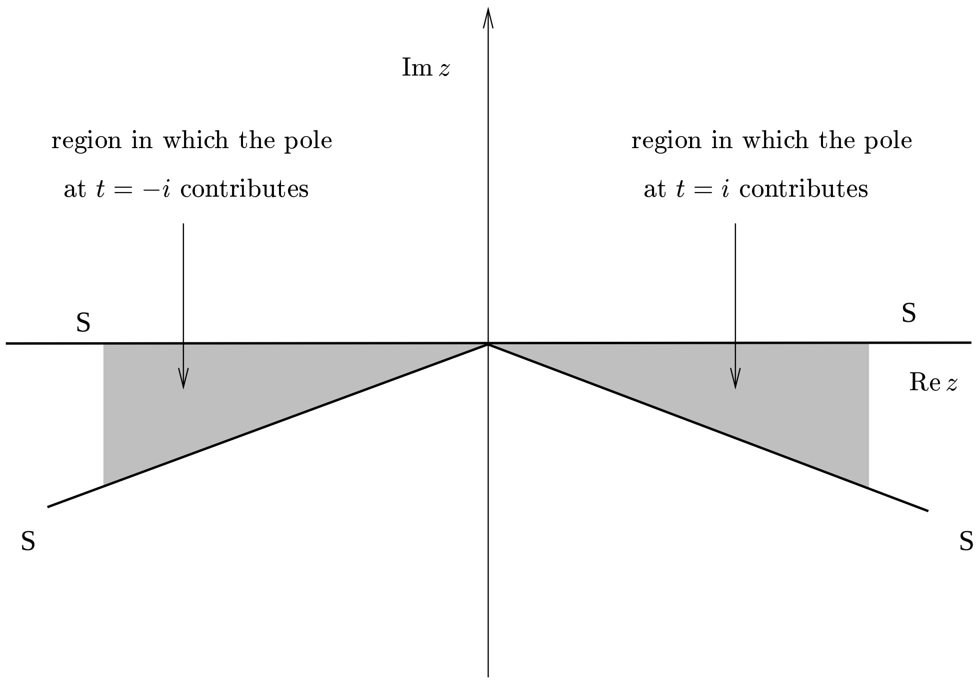


Figure 14: Schematic diagram showing the Stokes lines governing the contributions of the poles to the asymptotic behaviour of the integral (100).

We will henceforth define Stokes lines as lines across which a switching of the type described above actually occurs. The condition (98) necessarily holds on Stokes lines, but it is not sufficient to define them. However, this condition applies when integral representations may be unavailable; moreover both (98) and the steepest descent characterisations apply to linear partial differential equations and functions of several complex variables, in which case we will say the switching occurs across *Stokes surfaces*.

There is another characterisation of Stokes lines which does not demand prior representation of the solution in terms of an integral (such as a Fourier integral); we will discuss this in the next section.

4.3 Stokes' phenomenon via optimal truncation

The information about the "switching on" of a subdominant ray across a Stokes line is contained in the tail of the asymptotic expansion of the dominant ray [4, 34, 10]. We briefly review this discussion here.

Let us again use as an example the complementary error function,

$$\operatorname{erfc}(z) = \frac{2}{\sqrt{\pi}} \int_z^\infty e^{-v^2}\, dv,$$

which has the asymptotic expansion

$$\operatorname{erfc}(z) \sim \frac{e^{-z^2}}{z\sqrt{\pi}} \sum_{n=0}^{\infty} \frac{(-1)^n (2n)!}{n!(4z^2)^n} \qquad |\arg z| < 3\pi/4, \tag{101}$$

$$\operatorname{erfc}(z) \sim 2 + \frac{e^{-z^2}}{z\sqrt{\pi}} \sum_{n=0}^{\infty} \frac{(-1)^n (2n)!}{n!(4z^2)^n} \qquad |\arg(-z)| < 3\pi/4, \tag{102}$$

as $|z| \to \infty$. Now consider the expansion (101) in the sector $0 < \arg z < \pi$. We truncate after N terms to give

$$\operatorname{erfc}(z) = \frac{e^{-z^2}}{z\sqrt{\pi}} \sum_{n=0}^{N} \frac{(-1)^n (2n)!}{n!(4z^2)^n} + R_N(z), \tag{103}$$

from which it follows that the remainder $R_N(z)$ is a solution of the inhomogeneous ordinary differential equation

$$\frac{d^2 R_N}{dz^2} + 2z \frac{dR_N}{dz} = \frac{e^{-z^2}}{z\sqrt{\pi}} \frac{(-1)^N (2N)!}{(N-1)! 4^{N-1} z^{2N}}. \tag{104}$$

To examine the behaviour of R_N as $\arg z$ varies we introduce polar coordinates by writing

$$z = re^{i\theta}, \qquad 0 < \theta < \pi, \qquad \frac{d}{dz} = -\frac{ie^{-i\theta}}{r}\frac{d}{d\theta}, \qquad \frac{d^2}{dz^2} = -\frac{ie^{-2i\theta}}{r^2}\frac{d}{d\theta} - \frac{e^{-2i\theta}}{r^2}\frac{d^2}{d\theta^2}.$$

Then substitution into (104) gives

$$-\frac{e^{-2i\theta}}{r^2}\frac{d^2 R_N}{d\theta^2} + i\left(\frac{e^{-2i\theta}}{r^2} - 2\right)\frac{dR_N}{d\theta} = \frac{\exp\left[-r^2 e^{2i\theta} + N\pi i - (2N+1)i\theta\right](2N)!}{\sqrt{\pi}(N-1)! 4^{N-1} r^{2N+1}}. \tag{105}$$

Now suppose that we wish to truncate the expansion (101) for $\operatorname{erfc}(z)$ *optimally*, that is, we require the value of N for which R_N and hence the right-hand side of (105) is minimal. We find that this happens when $N \sim r^2$. We therefore set $N = r^2 + a$, where a is bounded as $r \to \infty$. Then, as $r \to \infty$ the right-hand side of (105) is approximately

$$\frac{8r}{\sqrt{2\pi}} \exp\left[-ia(2\theta - \pi) - i\theta - r^2(e^{2i\theta} + 1 + i(2\theta - \pi))\right]. \tag{106}$$

Following [34] we finally ask for which values of θ does the optimal remainder R_N change most rapidly, and note that (106) is maximal as a function of θ at the Stokes line $\theta = \pi/2$. At this point it also ceases to be oscillatory, and is independent of a. Introducing local coordinates near the Stokes line by setting

$$r = \rho/\epsilon, \qquad \theta = \pi/2 + \delta\psi,$$

we find

$$\frac{e^{-2i\delta\psi}\epsilon^2}{\rho^2\delta^2}\frac{d^2 R_N}{d\psi^2} - i\left(\frac{e^{-2i\delta\psi}\epsilon^2}{\rho^2\delta} + \frac{2}{\delta}\right)\frac{dR_N}{d\psi} \sim -\frac{8\rho i}{\epsilon\sqrt{2\pi}} e^{-2\rho^2\psi^2\delta^2/\epsilon^2}. \tag{107}$$

We find a sensible balance of terms when $\delta = \epsilon$ giving, to leading order

$$\frac{dR_N}{d\psi} \sim \frac{4\rho}{\sqrt{2\pi}} e^{-2\rho^2\psi^2}, \tag{108}$$

with solution

$$R_N \sim B + \operatorname{erf}(\sqrt{2}\rho\psi). \tag{109}$$

Matching (109) to the optimally truncated outer solution (101) as $\psi \to -\infty$ shows that the remainder is exponentially small, giving $B = 1$. As $\psi \to \infty$ we therefore find that $R_N \to 2$. Thus we can see in more detail the switching-on of the subdominant ray across the Stokes line.

We have now identified a further characterisation of the Stokes line $\theta = \pi/2$. In particular the right-hand side of (105) is maximal, non-oscillatory, and independent of a. These properties all follow from the fact that successive terms in the asymptotic expansion of $\operatorname{erfc}(z)$ have the same phase along the Stokes line [12]. This suggests the third characterisation in addition to those mentioned in §4.2, and this third characterisation also applies at once to functions of several complex variables; it is that surfaces on which the asymptotic expansion (1) of a general wavefield ϕ is divergent and satisfies

$$\lim_{n\to\infty} \arg \frac{iA_n}{A_{n-1}} = 0 \tag{110}$$

will be Stokes surfaces. This prescription gives us the possibility of proceeding *ab initio* from a WKB expansion of any particular wavefield and determining where other wavefields can be switched on and off, and thus it will prove extremely valuable in some of our later calculations.

We note that not all the Stokes lines or surfaces will be picked up this way. An example in [34] shows that the information about some Stokes lines may be buried deep in the expansion of A_n as $n \to \infty$. Consider for example the function $\operatorname{erfc}(z) + e^{-z^4}\operatorname{erfc}(iz)$. Information about the Stokes line of the second function on the positive real axis is swamped by the expansion of the first function there. Hence (110) will not select all the Stokes lines for such functions. However, when a sufficiently simple integral representation is not available, (110) can give extra information above (98) to help decide which of the equal-phase surfaces are Stokes surfaces.

In summary, we have two characterisations, (98) and (110), of Stokes' phenomenon, neither of which is complete, but which do not require an explicit representation of the solution in terms of an integral, and which can be used to identify Stokes surfaces as three-dimensional manifolds for ray theory in two real space dimensions. This will shortly enable us to re-examine our simplest motivational example, the circular caustic, with the aim of elucidating its asymptotic structure without prior knowledge of the formula (53). We will be especially interested in gaining *ab initio* information (i.e. just from (50)-(52)) about the position of the equal-phase surfaces and more importantly, which of these equal-phase surfaces are Stokes surfaces, that is to say actually switch on a subdominant rayfield. Hence, in §4.4, we will see how the complex rays are divided into active and inactive components by Stokes lines. This will intimately involve the Riemann surface of the multi-valued phase function, which is singular on the two-dimensional caustic manifold $x^2 + y^2 = 1$; of particular interest will be the way in which the boundaries of the Riemann surface relate to the Stokes surface.

4.4 The circular caustic revisited

(a) Inner region In §2.2 we derived the Airy equation

$$\frac{d^2 f}{d\rho^2} + 2\rho f = 0, \tag{111}$$

for the local field amplitude in the vicinity of the caustic. The conditions that we imposed on the solution to this equation were that in the interior of the caustic f is to be exponentially small (rather than large), whilst in the exterior two propagating rays must be present. Of these, precisely one should match the prescribed incoming rayfield, and the other is then consequently the outgoing rayfield. These conditions are enough to specify $f(\rho)$ uniquely, and it is proportional to $\mathrm{Ai}(-2^{1/3}\rho)$.

As a function of the complex variable $\rho = \rho_1 + i\rho_2$, $\rho_1, \rho_2 \in \mathbb{R}$, the Stokes structure of this solution is well known (see e.g. [35]), and is depicted in Fig. 15. The equal-phase lines are situated at $\arg\rho = \pi/3$, π, $5\pi/3$, whereas the anti-Stokes lines, along which the two asymptotic contributions are comparable, lie along $\arg\rho = 0$, $2\pi/3$, $4\pi/3$. Thus, outside the caustic in physical space ($\rho_2 = 0$, $\rho_1 \to \infty$) there are two rays of equal importance, and in the interior ($\rho_2 = 0$, $\rho_1 \to -\infty$) there is only one, exponentially small, contribution.

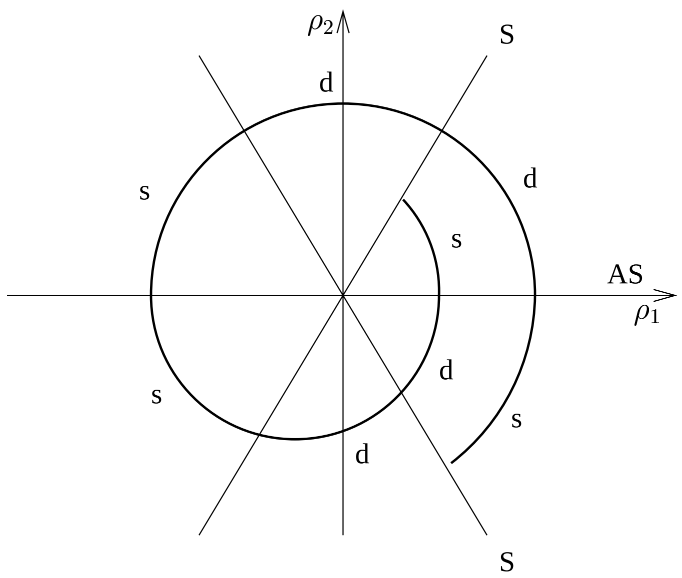


Figure 15: Diagram indicating the ray contributions in the WKB expansion of $\mathrm{Ai}(-2^{1/3}\rho)$; d and s indicate the dominant and subdominant rays respectively, while S and AS indicate the Stokes and anti-Stokes lines.

We emphasise that Stokes lines are associated with particular solutions of holomorphic differential equations, rather than the differential equations themselves. Of course, the existence of Fig. 15 is a consequence of prior knowledge of the solution of (111), say in [35]. For our purposes, it is important to ask whether it would be possible to deduce this information armed only with the governing equation (111), the matching conditions, and our knowledge of Stokes lines as discussed in §4.3. After all, a more general problem may not be as amenable as (111) to an exact analysis and even if it were, the solution might be so unwieldy as to make the construction of the Stokes line structure intractable. To show that in this case it is possible to deduce the Stokes line structure easily, we note that a WKB approximation to (111) as $|\rho| \to \infty$ provides the two ray solutions

$$f_-(\rho) \sim \frac{e^{-2\sqrt{2}i\rho^{3/2}/3}}{\rho^{1/4}}, \qquad f_+(\rho) \sim \frac{e^{2\sqrt{2}i\rho^{3/2}/3}}{\rho^{1/4}}.$$

Our "equal phases" argument (98) immediately limits the Stokes lines to $\arg\rho = \pi/3$, π, $5\pi/3$, and all that remains is to decide which of these are actually Stokes lines. We know that we need two contributing rays on $\rho_2 = 0$, $\rho_1 \to \infty$ and one (small) ray contribution on $\rho_2 = 0$, $\rho_1 \to -\infty$. Hence the equal-phase lines $\arg\rho = \pi/3$ and $\arg\rho = 5\pi/3$ must *both* be Stokes lines, and each must switch off one of the rays, since otherwise two rays and not just one would survive to reach $\rho_2 = 0$, $\rho_1 \to -\infty$. Thus, regardless of the direction in which we approach this half-axis, there can only be one ray contributing in its vicinity and so this equal-phase line *cannot* be a Stokes line. Of course, for other solutions of Airy's equation the equal-phase line $\arg(\rho) = \pi$ will be a Stokes line, and we shall see such an example in §5.

We note that since we have suppressed the coordinate θ in our inner analysis, the Stokes lines we have identified are really three-dimensional hypersurfaces in complex (x, y) space, and we will refer to them as Stokes surfaces henceforth.

This example suggests the following procedure: determine the possible contributing complex rayfields, determine the equal-phase surfaces associated with them taken in appropriate pairs, and then decide which of these are Stokes surfaces. This should then provide a systematic recipe for finding "regions of existence" of complex rays and, most importantly, the implications that this has for the physical solution. The most difficult part of this procedure in general will be deciding which of the equal-phase surfaces are "active" Stokes surfaces. In some cases, as in the present example, this can be seen immediately from inspection. However, in other cases the Stokes surfaces are not immediately obvious, and we may need to call on the other characterisation of Stokes surfaces that we have identified earlier in this section, namely (110).

(b) Outer region Well away from the caustic, the ray solution (50)-(52) is valid. The Stokes surfaces emanating from the caustic can be continued into the outer region using the first of our characterisations, namely, as surfaces of equal phase (as noted by [16] for the case of a circular caustic of complex radius). From (56) the equation for the equal-phase surfaces is

$$\mathrm{Re}\left[\cos^{-1}(1/r) - \sqrt{r^2 - 1}\right] = 0. \tag{112}$$

These surfaces are therefore no longer flat and whether or not the surfaces represented by (112) are Stokes surfaces is determined by matching with those in Fig. 15 as the caustic is approached. Two different cross-sections are shown in Figs. 16 and 17. Consider the first of these, which is a planar cross-section for $y = 1/2$. This intersects the caustic at the real points $(\pm\sqrt{3}/2, 1/2)$, each of which therefore has a behaviour as in Fig. 15 in its vicinity. To understand the implications of Fig. 16, restrict attention to the local structure near $(-\sqrt{3}/2, 1/2)$. We know that in $x_2 = 0$, $x_1 > -\sqrt{3}/2$ (i.e. inside the caustic) there is one complex ray yielding an exponentially small solution. This, of course, is guaranteed for us by the Stokes structure of the inner Airy solution. As we sweep anti-clockwise around this inner region we meet the first solution of (112), which is a Stokes surface extending to infinity in the second quadrant (shown as the continuous curve). Since it must match with the inner Stokes surface, we see that it is active and it switches on a subdominant complex ray. This then grows in amplitude whilst the original diminishes until they are of a comparable size along the anti-Stokes surface $x_2 = 0$, $x_1 < -\sqrt{3}/2$. This, of course, provides the two propagating rays exterior to the caustic. As we sweep still further, we meet another solution of (112), this time shown as a dotted curve in the third quadrant. This is also active, again because of matching in to the Airy Stokes structure, and switches off the original ray with which we started which is now subdominant when compared to the new complex ray switched on at the previous Stokes line. The one remaining complex ray now survives all the way back to $x_2 = 0$, $x_1 > -\sqrt{3}/2$. The solution it yields there is identical to that given by the original ray. Note that when this ray was switched on at the first Stokes line, it corresponded to the branch of the square root in (56) different from the ray with which we started. What has happened is that we have now circled the branch point, and the two ray solutions are actually two sheets of the Riemann surface of the phase function u. If we wish to assign a "label" to each ray at each point in complex space (such as the "incoming" and "outgoing" labels which distinguish between the two rays in $x_2 = 0$, $x_1 < -\sqrt{3}/2$), then as we circle the branch point our label will have to switch at some point (the branch cut). We prefer, however, to work with the Riemann surface, rather than insert branch cuts, and label the rays only when they are real. In §6 we will find the need to label rays for all complex x and y, and in that case we will have to introduce branch cuts. We shall return to this point shortly when we discuss the detailed structure within an individual complex ray.

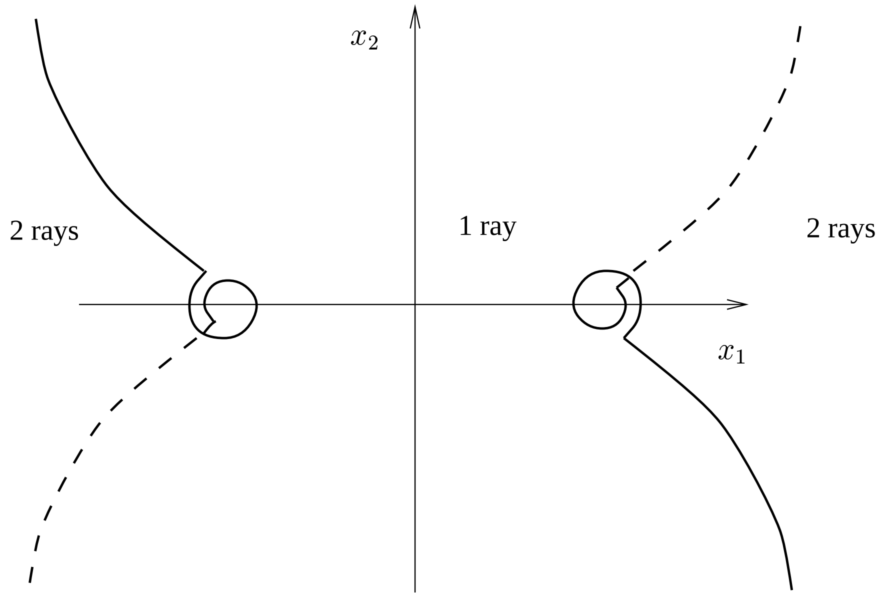


Figure 16: Diagram showing the Stokes surfaces for the circular caustic in the plane $x = x_1 + ix_2$, $y = 0.5$, with x_1, $x_2 \in \mathbb{R}$. The dotted line indicates the cut in the lower sheet of the Riemann surface.

Fig 17 shows a different cross-section, this time for $y = 2$. In this case the intersection with the caustic is at $(\pm i\sqrt{3}, 2)$ and the Stokes surfaces must therefore emanate from the vicinities of these points. Notice that in this case, the intersection of the cross-section with the physical domain $x_2 = 0$, $-\infty < x_1 < \infty$ contains two propagating real rays, as expected.

We now return to the question of the structure within any one ray. The way we have set things up, the real rays meet the caustic at $t = 0$, so that negative and positive values of t correspond to incoming and outgoing portions of the ray, and this split is unambiguous for real values of t. We might then ask about corresponding interpretations, if there are any, for complex values of t. More importantly, we need to

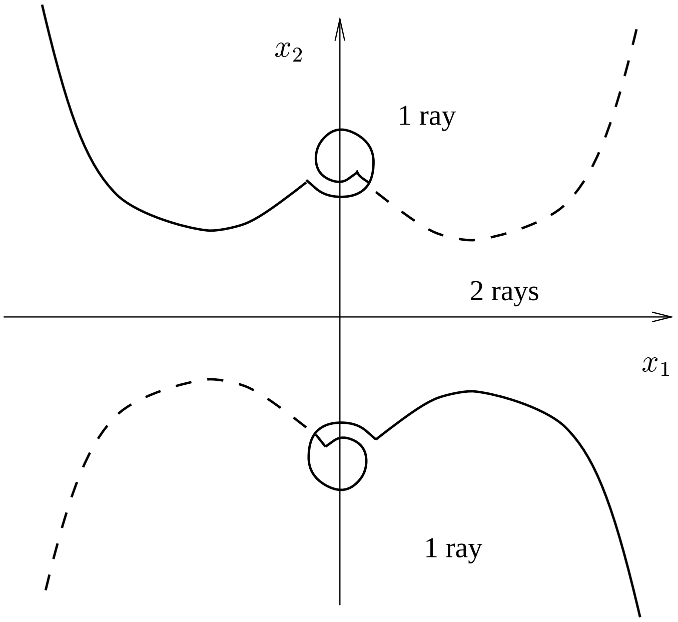


Figure 17: Diagram showing the Stokes surfaces for the circular caustic in the plane $x = x_1 + ix_2$, $y = 2$, with x_1, $x_2 \in \mathbb{R}$. The dotted line indicates the cut in the lower sheet of the Riemann surface.

know whether or not the ray actually exists for all complex values of t; after all, we have seen that complex rayfields do not exist everywhere since they are switched on and off across Stokes surfaces. Thus it is not unreasonable to expect that individual complex rays are themselves "cut" along certain directions. [4]

We begin by noting that by symmetry the structure within all rays will be the same, although in more general circumstances this will not necessarily be the case. Using the ray equations (50)-(52) to write

$$t = (r^2 - 1)^{1/2}, \tag{113}$$

the equal-phase condition (112) can be written in the form

$$\mathrm{Re}\left[\tan^{-1} t - t\right] = 0, \tag{114}$$

and this is shown graphically in Fig. 18. This figure shows three equal-phase curves, BE, AD and CF. We observe that the branch point behaviour of the phase u as a function of r has been "opened out" by considering u as a function of t (cf. (113)) instead. The individual portions A and D of AD map to the same line in r space; on any individual ray, the lines A and D correspond to different values of θ, but for each point on the line A on one ray there is another ray for which the corresponding point on the line D maps to the same point in (x, y)-space.

For values of t on the lines B, D and F the ray is the dominant ray of the two rays through that point in space, while for values of t on the lines A, C and E the ray is subdominant. The equal-phase lines B and E

[4] The fact that this can happen in practice is illustrated by the excitation of a surface ray on an elastic structure coupled with a compressible fluid. In [39] it was found that the surface ray was excited by a critically-incident complex ray, but that it did not occupy all regions of the boundary. This is an example of a complex ray (the surface ray) lying in a two-dimensional, planar manifold but only occupying a portion of this plane. The ray is cut across a certain direction and only exists on one side of it. The analysis in [39] gave a strong indication that the mechanism for doing this is Stokes' phenomenon.

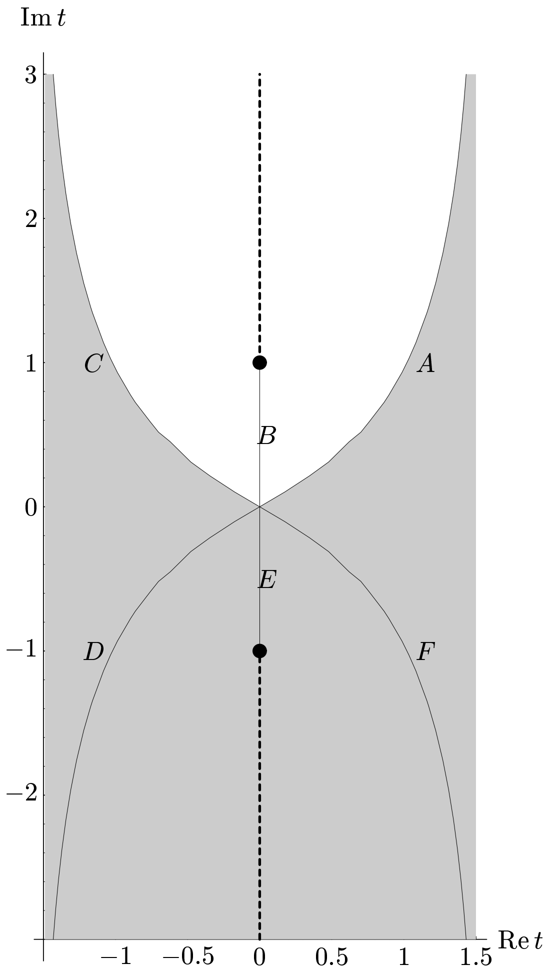


Figure 18: Structure of a single ray ($t \in \mathbb{C}$, s constant) for the circular caustic.

correspond to the equal-phase line which is not a Stokes line. On the equal-phase line A the ray is switched off by the second, dominant, ray, whose value of t lies on the line D. Similarly, on the equal-phase line C the ray is switched off by the second ray, whose value of t lies on the line F.

Thus the diagram shows that the ray is “switched off” across the two curves A and C, and only exists for complex values of t lying below these two curves. Where this switching occurs can, of course, also be interpreted as the intersection of the global Stokes surface engendered by the caustic with the complex ray. Each complex ray will either intersect the caustic at a real point, in which case its intersection with real space is a propagating ray, or at a complex point, in which case it will pass though the interior of the caustic. Either way, the structure of the ray is the same. Suppose the ray is of the former type. Values of t for which $t_2 = 0$, $t_1 > 0$ ($t_1 < 0$) can then be identified as the outward (inward) propagating portion of a real ray having passed the tangency point $t = 0$ with the caustic. This ties in with the ideas of Fig. 16 where we can see that in the region given by $x_2 = 0$, $x_1 > \sqrt{3}/2$ there are two rays, one incoming and the other outgoing. Inside the caustic in real space, where $x_2 = 0$, $x_1 < \sqrt{3}/2$ in Fig. 16 we see that there is only one ray. Since $r < 1$ in this case, t must be pure imaginary, and since Fig. 18 shows us that no ray contains the range of values of t for which $t_1 = 0$, $t_2 > 0$ we deduce immediately that this ray must be “initiated” at a complex point on the caustic and propagate a negative pure imaginary “distance” to reach the interior of the physical caustic.

Thus the planes which are the rays are also cut along Stokes surfaces. The Stokes surfaces determine which parts of the ray are active, and in this sense they replace the laws of causality applied to real rays.

(c) Optimal truncation Having identified the Stokes surfaces for the circular caustic, let us see how they arise via optimal truncation of the dominant ray expansion. This will shed further insight into the mechanism which decides which of the equal-phase surfaces are Stokes surfaces.

For the circular caustic solution of the transport equations we have

$$A_0 = \frac{1}{\sqrt{-t}}, \qquad A_n \text{ independent of } s \text{ for all } n.$$

Hence, using the general expression (30) we find

$$A_n = \sum_{j=n}^{3n} \frac{g_{n,j}}{t^{j+1/2}}, \tag{115}$$

where

$$g_{n+1,j} = -\frac{1}{2j}\left[(j-1/2)^2 g_{n,j-1} + (j-5/2)(j-1/2)g_{n,j-3}\right], \tag{116}$$

with $g_{0,0} = i$. We expect that $g_{n,j}$ will be of the order of $n!$ as $n \to \infty$, so we define $g_{n,j} = (-1)^n n! b_{n,j}$. We can also reduce the order of the recurrence relation by setting

$$k = \frac{j-n}{2}, \quad b_{n,j} = c_{n,k}, \quad b_{n,j+2} = c_{n,k+1}, \quad b_{n+1,j+1} = c_{n+1,k},$$

to give

$$c_{n+1,k+1} = \frac{(2k+n+5/2)^2}{2(n+1)(2k+n+3)} c_{n,k+1} + \frac{(2k+n+5/2)(2k+n+1/2)}{2(n+1)(2k+n+3)} c_{n,k}. \tag{117}$$

Taking the dominant balance as $n \to \infty$ gives

$$c_{n+1,k+1} \sim \left(\frac{2k+n}{2n}\right)\left[c_{n,k+1} + c_{n,k}\right]. \tag{118}$$

Now setting $w = k/n$, $c_{n,k} \sim \hat{B}(n,k)e^{n\psi(w)}$, where $|\log(\hat{B})| = o(n)$, and letting $n \to \infty$ gives

$$e^{\psi}e^{(1-w)\psi'} = \left(w + \frac{1}{2}\right)\left(1 + e^{\psi'}\right), \tag{119}$$

where a prime denotes d/dw. We can solve for $g_{n,n}$ and $g_{n,3n}$ exactly to give the following boundary conditions on (119):

$$\psi(0) = \log 1/2, \tag{120}$$
$$\psi(1) = \log 3/2. \tag{121}$$

Note that since (119) is first-order we require only one boundary condition; the other should be automatically satisfied, and provides a check on the analysis.

Let us return to the expression for the A_n for a moment. We have

$$\begin{aligned} A_n &= \sum_{j=n}^{3n} \frac{g_{n,j}}{t^{j+1/2}} \\ &= \sum_{j=n}^{3n} \frac{(-1)^n n! b_{n,j}}{t^{j+1/2}} \\ &\sim \frac{n!}{\sqrt{t}} \sum_{j=n}^{3n} \frac{(-1)^n B(n,j) e^{n\psi((j-n)/2n)}}{t^j}, \end{aligned} \tag{122}$$

where $B(n,j) = \hat{B}(n,k)$. Writing $j = vn$ and approximating the final sum by an integral gives

$$A_n \sim \frac{nn!}{t^{1/2}} \int_1^3 B(n,vn) \exp\left(n\left[\psi\left(\frac{v-1}{2}\right) - v\log t + i\pi\right]\right) dv. \tag{123}$$

Defining

$$f(v) = \psi\left(\frac{v-1}{2}\right) - v\log t,$$

we have

$$A_n \sim \frac{nn!}{t^{1/2}} \int_1^3 B(n,vn) \exp\left(n\left[f(v) + i\pi\right]\right) dv, \tag{124}$$

with

$$t^3 e^f e^{(3-v)f'} = \frac{v}{2}\left(1 + t^2 e^{2f'}\right), \tag{125}$$

where a prime represents d/dv. Applying the method of steepest descents to (124) we find the dominant contribution occurs near v_0 where $f'(v_0) = 0$, and is given by

$$A_n \sim CB(n, v_0 n)\frac{n\,n!}{t^{1/2}} \exp\left(n\left[f(v_0) + i\pi\right]\right) \tag{126}$$

for some constant C. From (125) we see

$$e^{f(v_0)} = \frac{v_0(1+t^2)}{2t^3}. \tag{127}$$

Thus we need to find v_0. Differentiating (125), using (125) to eliminate f, and writing $g = f'$ gives

$$g' = \frac{1 + t^2 e^{2g}}{v(t^2 e^{2g}(1-v) + 3 - v)}, \tag{128}$$

with boundary conditions $g(1) = \infty$, $g(3) = -\infty$. By writing $v = v(g)$, we find the solution to (128) is given by

$$v = \frac{t^3 e^{3g}}{(1 + t^2 e^{2g})(te^g - \tan^{-1}(te^g) + a)}. \tag{129}$$

The boundary conditions imply that $a = 0$.

From (126) and (127) we see

$$A_n \sim CB(n, v_0 n) \frac{n\, n!}{\sqrt{t}} \left(\frac{-1}{2(t - \tan^{-1} t)} \right)^n. \tag{130}$$

Now that we have the leading order behaviour of A_n, we can optimally truncate and observe the Stokes switching exactly as in §4.3. The condition (110) gives the Stokes surfaces as

$$\frac{-i}{2(t - \tan^{-1} t)} \qquad \text{real } > 0. \tag{131}$$

Note that this is consistent with the equal-phase surfaces (114), namely

$$\mathrm{Re}[\tan^{-1} t - t] = 0, \tag{132}$$

but it also identifies which of the equal-phase surfaces are Stokes surfaces. As $t \to 0$, $t - \tan^{-1} t \sim t^3/3$. Thus the Stokes surfaces are given by $\arg t = \pi/2$, $7\pi/6$, $11\pi/6$, as $t \to 0$. The ray is not present at $\arg t = \pi/2$, and so this line cannot be a Stokes surface. The other two are exactly the lines D and F of Fig. 18.

Finally, we note that there may be an easier way to deduce the leading-order asymptotic behaviour of A_n than the one employed. Anticipating replacing the sum in the expression for A_n by an integral and performing steepest descents, we expect to find

$$A_n \sim \frac{\Lambda \Gamma(n + \gamma)}{F^n} \text{ as } n \to \infty, \tag{133}$$

where Λ, γ and F may be functions of s and t but are independent of n (in (126) $F = e^{i\pi + f(v_0)}$). If we use this ansatz directly in (25) and retain only leading order terms as $n \to \infty$ we find

$$-2F_\tau = \alpha^2 F_\sigma^2 + F_\tau^2 + \frac{2\alpha^2(\beta' - u_0')F_\sigma F_\tau}{\tau^2} + \frac{\alpha^2(\beta' - u_0')^2 F_\tau^2}{\tau^2}. \tag{134}$$

In the present example $\beta = 0$, $\alpha = 1$, $u_0' = 1$, and $F = F(\tau)$, giving

$$F_\tau = -\frac{2\tau^2}{1 + \tau^2}. \tag{135}$$

Hence

$$F = -2(\tau - \tan^{-1} \tau)$$

in agreement with (130).

Finally a word of warning when using (134) to determine the behaviour of A_n as $n \to \infty$. The ansatz (133) is very similar to a WKB ansatz, and exactly the same problems associated with analytic continuation and Stokes lines occur in the approximation of A_n as in the original approximation of ϕ (this is related to the phenomenon known as resurgence [5]). We can see this more clearly with reference to (124). In general f will have many saddles, and, as s and t vary, different saddles may be picked up by the integral, other saddles may dominate, and the saddle we started with may be switched off. Thus care needs to be taken, and this limits the usefulness of (134).

5 Non-analytic solutions of Helmholtz equation

In solving the inner problem for the circular caustic, equation (63), we chose the solution which was exponentially small as $\rho \to -\infty$, giving $f = C\,\mathrm{Ai}(-2^{1/3}\rho)$. Let us consider here two other solutions of (63), given by choosing different branches of the cube root of 2, namely

$$f = C\,\mathrm{Ai}(-e^{2i\pi/3}2^{1/3}\rho), \tag{136}$$

$$f = C\,\mathrm{Ai}(-e^{4i\pi/3}2^{1/3}\rho). \tag{137}$$

The Stokes lines for these solutions are given by rotating Fig. 15 through $2\pi/3$ and $4\pi/3$ radians respectively, and are shown in Fig. 19.

In each case there are two rays present as $\rho \to -\infty$, and so the solution is exponentially large there, but only one ray is present as $\rho \to \infty$. The solution (136) shown in Fig. 19(i) has the outgoing ray only as $\rho \to \infty$, while solution (137) shown in Fig. 19(ii) has the incoming ray only, as shown in Fig. 20.

The solution (136) is the caustic for the acoustic ring source [9], and is produced, for example, by a small propeller rotating at a Mach number less than one in the interior of the circle $r = 1$. The caustic is at the sonic radius, and the outgoing wavefield is exponentially small by comparison with the wavefield in the vicinity of the propeller.

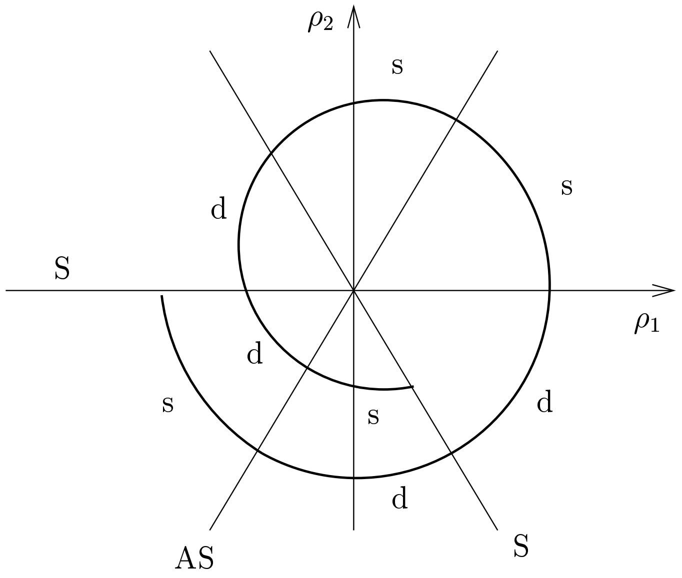
When we examine the Stokes surface structure of the outer problem for the acoustic ring source, we find we have a different set of active Stokes surfaces from those shown in Figs. 16 and 17, and the corresponding picture is shown in Figs. 21 and 22. The Stokes surface cuts the Riemann surface of u in such a way that the incoming wavefield is not present in the exterior of the caustic. However, we see from Figs. 21 and 22 that another problem has emerged. The Stokes surface on the imaginary axis terminates at the points $x = \pm iy$, corresponding to $r = 0$. What is then to stop us analytically continuing the solution around this point to obtain the physically unrequired incoming wavefield in $x^2 + y^2 > 1$, as indicated in Fig. 23?

If we examine the rays for this solution we see the same problem. The active portion of each ray is shown in Fig. 24. The Stokes surface meeting the ray on $t_1 = 0$ again terminates at $t = i$, corresponding to $r = 0$; the Stokes surface itself does not divide the ray into two halves.

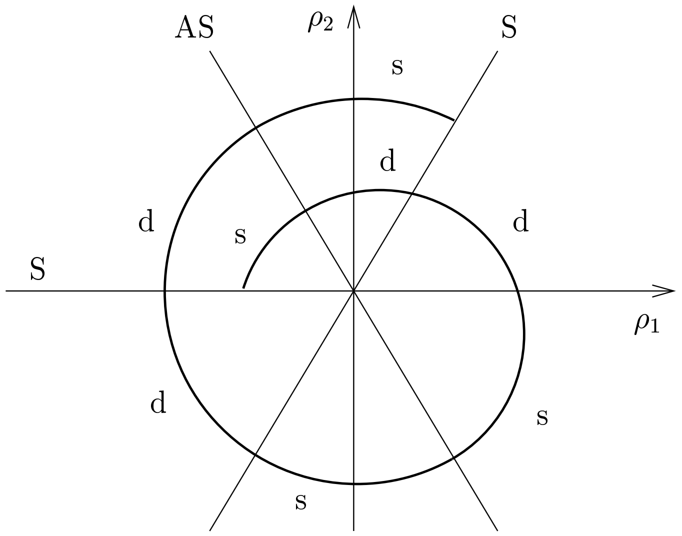
The root of this problem becomes immediately clear if we go back to consider the exact solution of Helmholtz equation. Separating the variables in polar coordinates gives $\phi = f(r)e^{ik\theta}$ with $f = J_k(kr) + iY_k(kr)$. Hence in this case f *itself* has a branch point at the origin and is in fact a multi-valued function living on a Riemann surface with infinitely many sheets. Thus when we analytically continue around $r = 0$ we pass onto a different sheet of the Riemann surface. Hence we have to include branch cuts in Figs. 21 and 22 to indicate the end of one Riemann sheet and the start of the next, as shown in Figs. 25 and 26.

There is an important point to be made here. We have emphasised earlier that the phase u is a multi-valued function living on a Riemann surface, and that this surface ends, or is cut, at the Stokes surfaces of the solution; the branch points of u are the caustics of the wavefield. For the acoustic ring source, ϕ *itself* is a multi-valued function, whose branch cuts we have had to insert as part of our solution procedure in order to ensure single-valuedness.

In the first case the branch points are branch points of the *approximation*, and are smoothed via an inner analysis in the vicinity of the caustic. There is no singularity in the exact solution, and the “cuts” in the Riemann surface are not arbitrary, but are determined by Stokes’ phenomenon. The cuts do not indicate

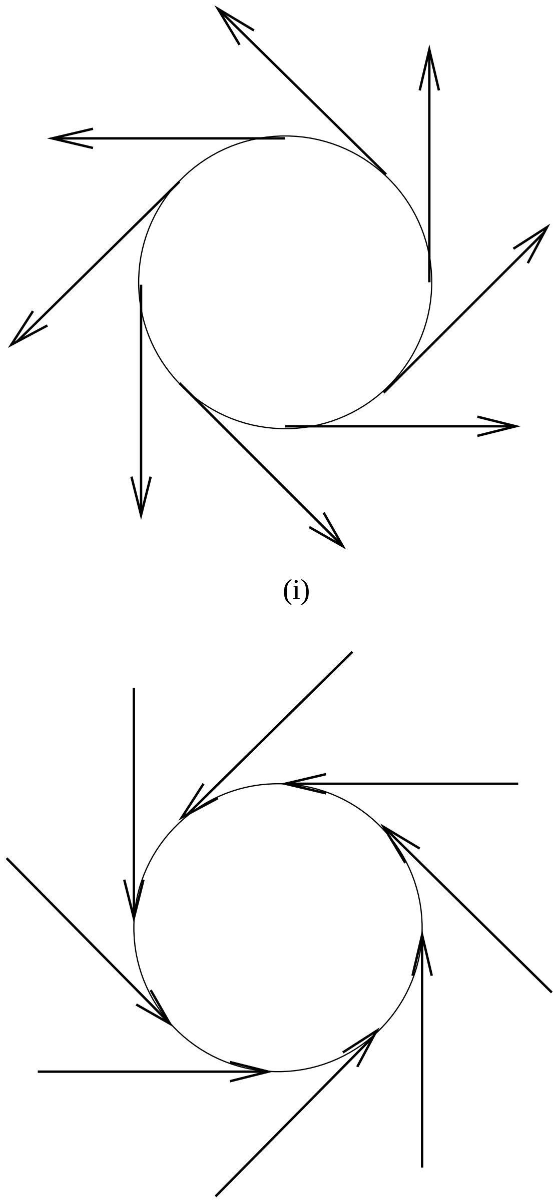


(a) structure of (136)



(b) structure of (137)

Figure 19: Stokes lines and ray structure for the solutions (i) (136) and (ii) (137); d and s indicate the dominant and subdominant rays respectively, while S and AS indicate the Stokes lines and anti-Stokes lines.



(i)

(ii)

Figure 20: Rays contained in the far-field approximations of (136) and (137).

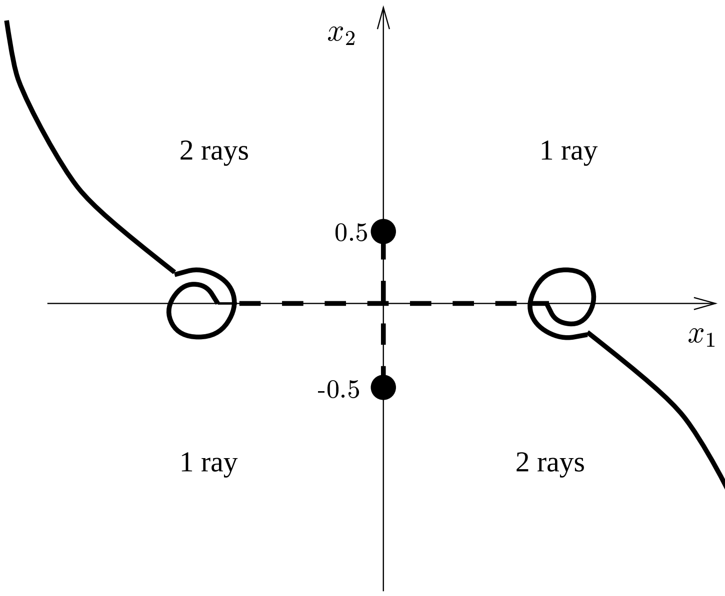


Figure 21: Stokes surface for the acoustic ring source in the plane $x = x_1 + ix_2$, $y = 0.5$; x_1, $x_2 \in \mathbb{R}$.

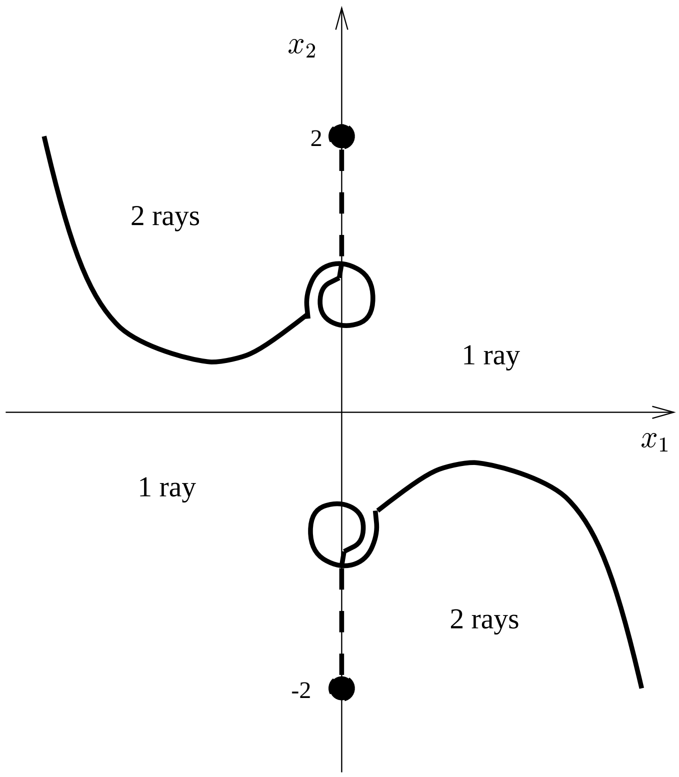


Figure 22: Stokes surface for the acoustic ring source in the plane $x = x_1 + ix_2$, $y = 2$; x_1, $x_2 \in \mathbb{R}$.

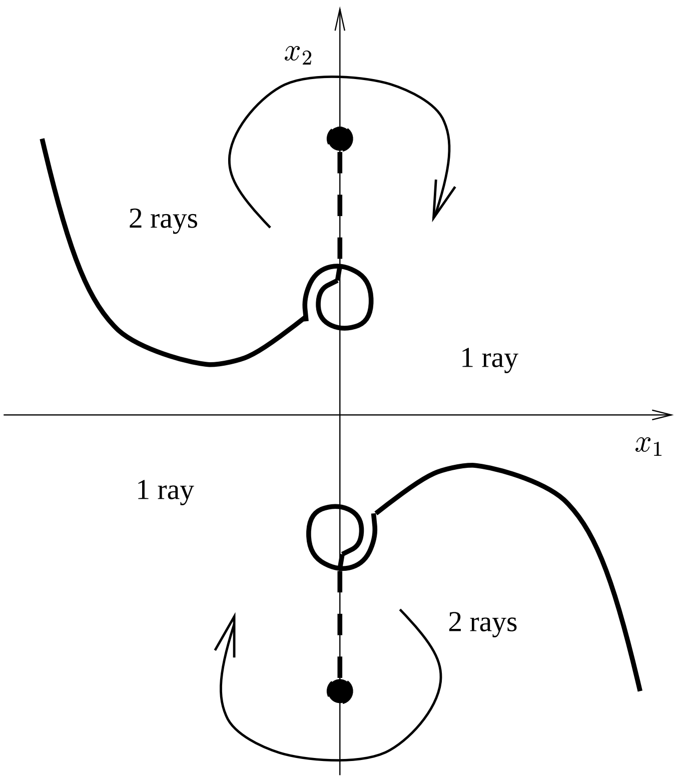


Figure 23: Analytically continuing to the other side of the Stokes surface in Fig. 22.

that we have moved onto another sheet of the Riemann surface, rather that that sheet is not present (i.e. that the amplitude of the corresponding wavefield is zero). For example, in the region labelled “2 rays” in Fig. 16 the rays corresponding to both sheets of the Riemann surface of u are present, but in the region labelled “1 ray” the ray corresponding to the one sheet of the Riemann surface is not present — it was switched off across the Stokes surface.

In the second case the singularity *is* in the exact solution, so that the exact solution is multi-valued and we must use radiation/extinction conditions to decide which solution branch is the physically relevant one. The branch cuts may be placed arbitrarily, and do not represent an end to the Riemann surface, rather a jump from one sheet to another.

If we now move to look at the ray picture, we find that the rays themselves are Riemann surfaces, with branch points at $t = \pm i$, and that branch cuts must be inserted in Fig. 24 also, as shown in Fig. 27. There is now no contradiction regarding the active part of each ray.

We see that in fact the outgoing and incoming caustics (136) and (137) each correspond to one sheet of the multi-valued solution, that is, they are joined together by analytic continuation. The other sheets correspond to multiples of the circular caustic solution superimposed on either of these.

In summary, when the exact solution has singularities, the Stokes surfaces may terminate and branch cuts will need to be introduced into the Riemann surface of the phase and the rays. As we discussed in §3, such situations are generic, since when we analytically continue, say, inside a closed reflecting body we are effectively solving the Cauchy problem in a bounded domain. In fact, the “outgoing-only” caustic can

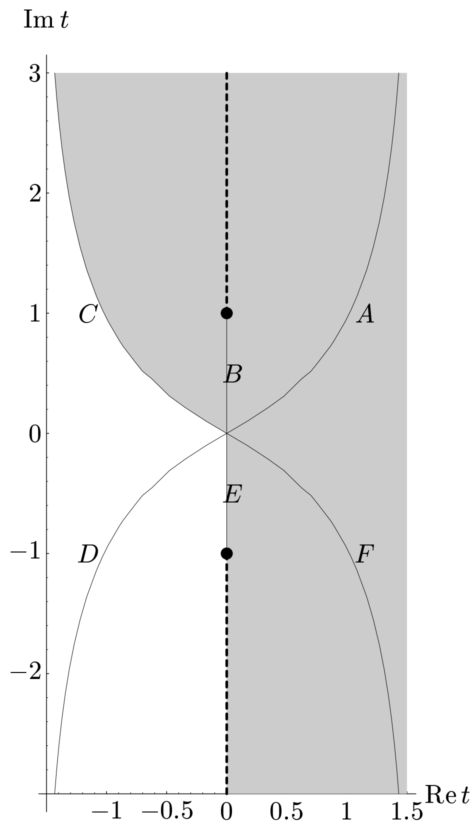


Figure 24: Ray structure for the Acoustic Ring Source.

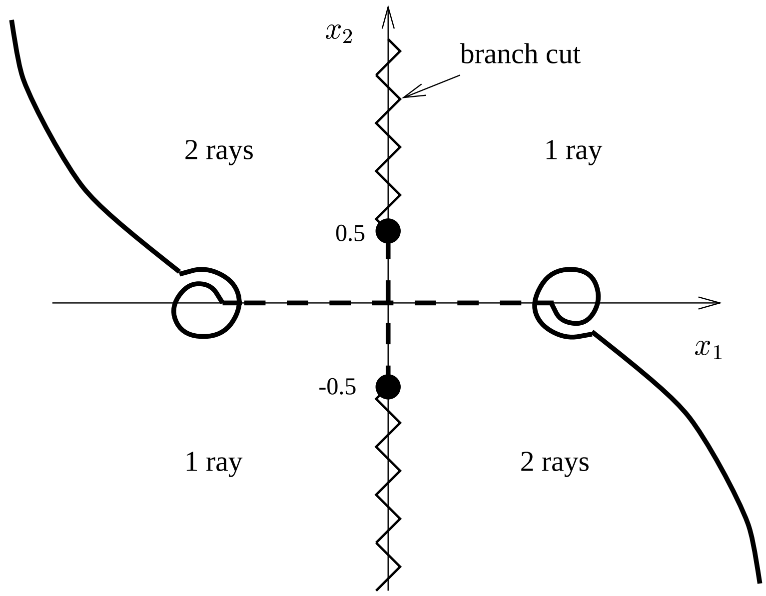


Figure 25: Stokes surface and branch cuts for the acoustic ring source in the plane $x = x_1 + ix_2$, $y = 0.5$; x_1, $x_2 \in \mathbb{R}$.

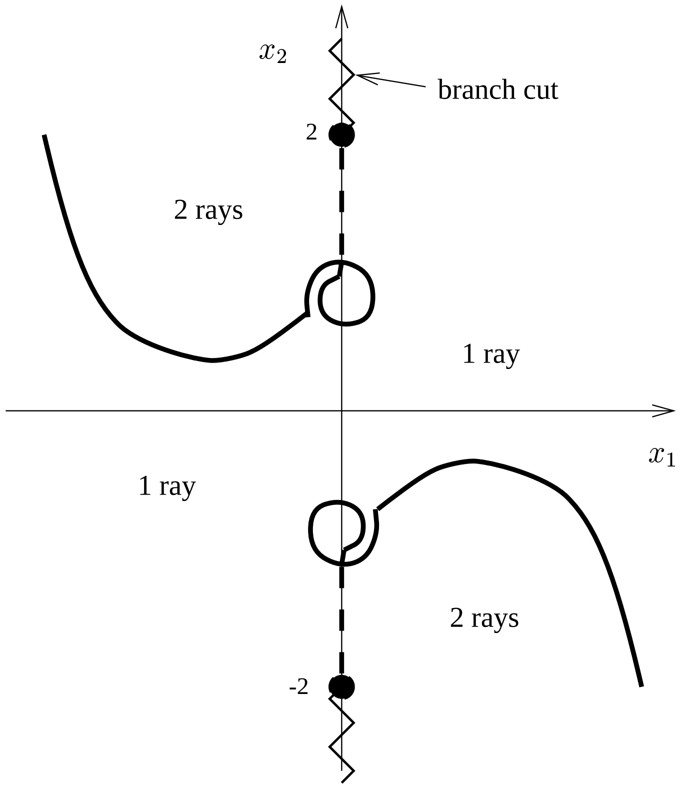


Figure 26: Stokes surface and branch cuts for the acoustic ring source in the plane $x = x_1 + ix_2$, $y = 2$; $x_1, x_2 \in \mathbb{R}$.

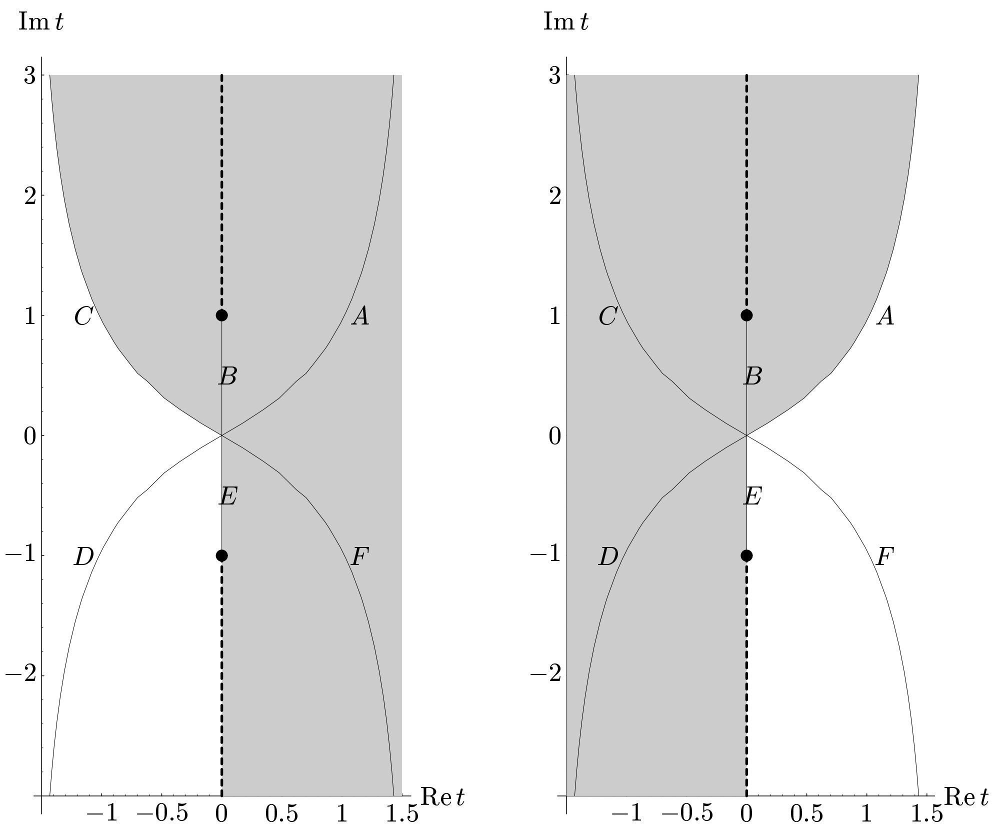


Figure 27: Ray structure for the Acoustic Ring Source (shaded region shows where the ray is active).

account for some of the spurious rays in Fig. 11. If we analytically continue the reflected field back into a convex scatterer there will be a caustic of the wavefield in the interior (see Fig. 28). Since we do not want the incoming part of the rays to be present for the solution to be physically acceptable, this caustic must be an "outgoing-only" caustic. The solution inside the caustic will be exponentially large, and may well contain branch points, but this will not happen in physical space.

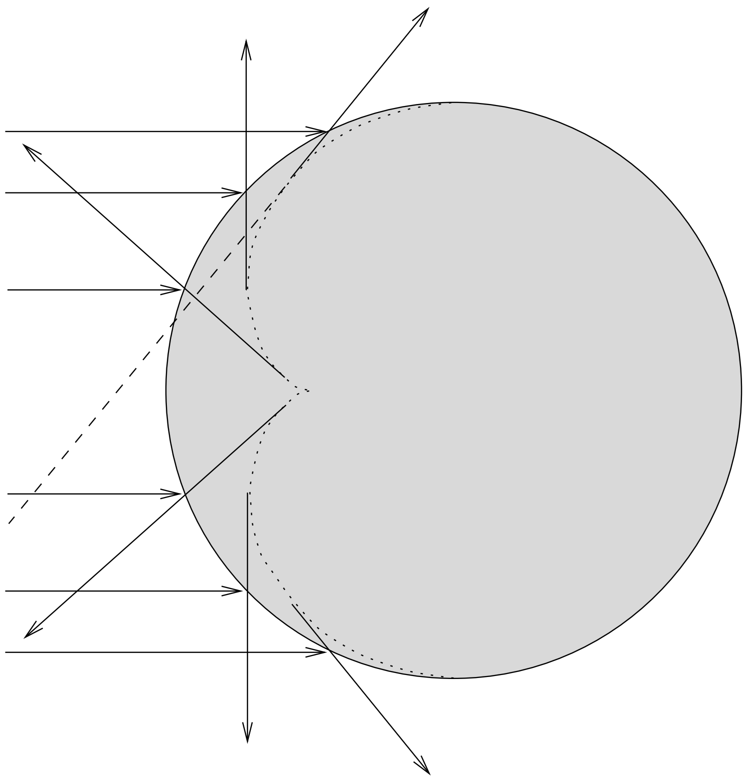


Figure 28: Caustic (dotted curve) obtained by analytically continuing the reflected field due to an incident plane wave on a circular cylinder back into the scatterer.

6 Examples

Even though we do not have systematic rules for determining which of the equal-phase surfaces are Stokes surfaces, let us now illustrate our theory of complex rays by applying it to some specific problems.

We will consider first a model problem, which has similarities with the problem of total internal reflection, but which has the advantage that there is a relatively simple Fourier transform solution.

6.1 Model problem

Consider the following model problem:

$$\nabla^2\phi + k^2\phi = 0 \qquad \text{in } y > 0, \tag{138}$$

$$\phi = e^{ikx^2/2} \qquad \text{on } y = 0, \tag{139}$$

$$\phi \qquad \text{outgoing at infinity.} \tag{140}$$

Solving the ray equations we find that $p = s$, $q = \sqrt{1-s^2}$, so that

$$\begin{aligned} x &= (1+t)s, & (141) \\ y &= t\sqrt{1-s^2}, & (142) \\ u &= s^2/2 + t, & (143) \end{aligned}$$

where the positive branch of the square root is taken for $-1 < s < 1$, so that the real rays are outgoing. The real ray solution is illustrated in Fig. 29.

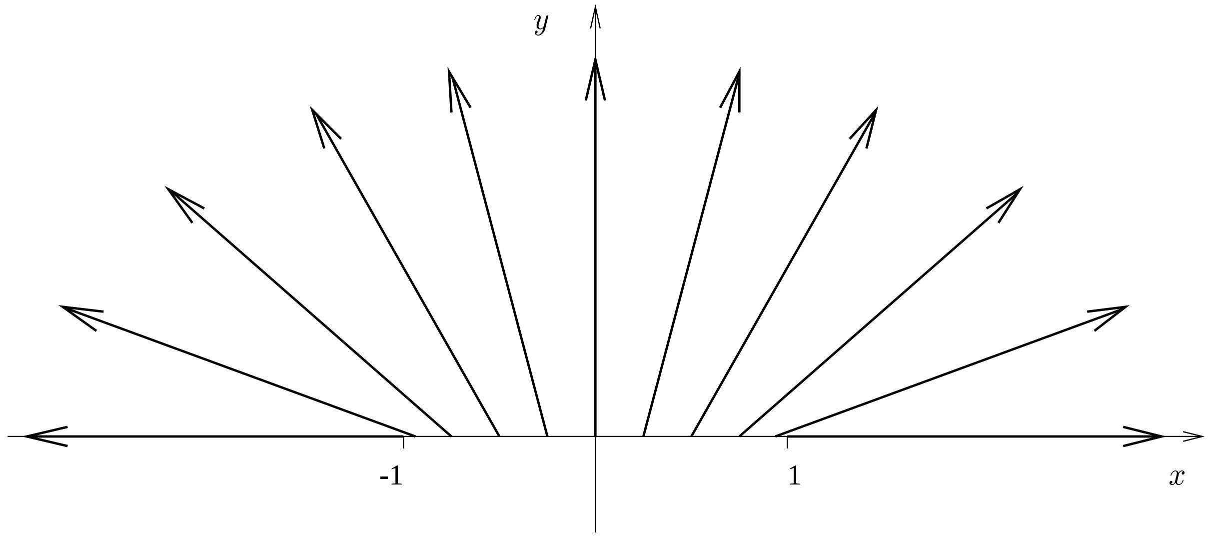


Figure 29: Real rays in the problem (138)-(140)

There are two points to notice. Firstly, rays coming from points on the boundary $|s| > 1$ immediately leave real space. Secondly, the real ray solution (141)-(143) restricted to $|s| \leq 1$ does not satisfy the boundary data on $|x| > 1$. Thus there must be an evanescent wavefield present for $|x| > 1$, $y \ll 1$, corresponding to complex rays. Let us use our theory of complex rays to calculate this wavefield and its domain, that is, the region in which complex rays should be added to obtain exponential accuracy in the solution.

We first determine the number of rays present when (141), (142) are complexified. Eliminating t gives

$$(x-s)^2(1-s^2) = s^2y^2. \tag{144}$$

This is a quartic for s, and hence there will be four rays through any one point. We can identify these rays more easily by examining the caustics associated with the wavefield. We find from (14) that $\beta(s) = 1 - s^2$, so that the caustic is given by $t + 1 - s^2 = 0$, or equivalently

$$x^{2/3} + y^{2/3} = 1.$$

This curve is plotted in real space in Fig. 30.

The rays can now be identified as the tangents to this curve. Through points inside the caustic we see that there are four real tangents, so that all the rays are real. As we cross the caustic two points of tangency

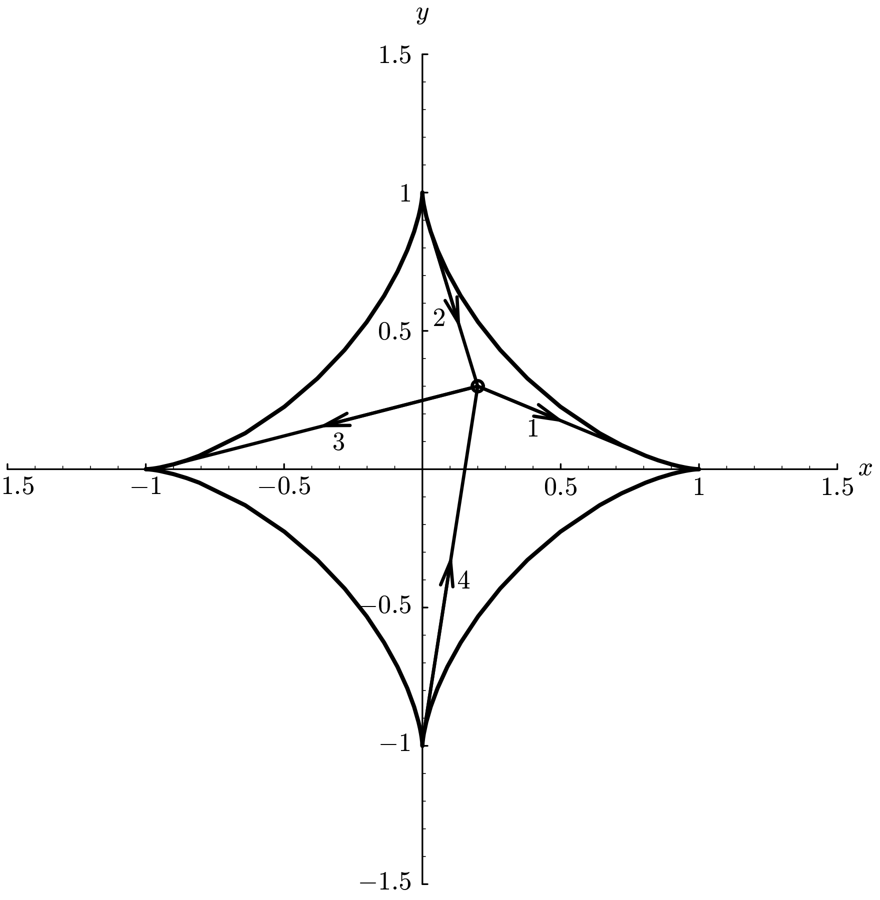


Figure 30: Caustic of the rays (141)-(142)

coalesce, and two of the real rays become complex rays. Thus, through points outside the caustic there are two real rays and two complex rays.

Of the four real rays, ray number 4 corresponds to the required outgoing ray, while rays 1, 2 and 3 correspond to incoming rays and should therefore not be present. Of the two complex rays one will form our evanescent wavefield, but the other will be exponentially large and not present.

We know that the complex ray must be present for $|x| > 1$, $y \ll 1$. However, we know that it cannot be present inside or near the caustic in the upper half plane, since the real rays that form this caustic are incoming and therefore not present. Thus the complex ray must be switched off across some lines in the upper half place (the intersection with the Stokes surfaces), and these lines must pass through the points $(-1, 0)$ and $(1, 0)$ respectively.

To identify the Stokes surfaces we first need to find the equal-phase surfaces, that is, all surfaces on which two of the four rays have equal phase. We find that the two complex rays have equal phase everywhere outside the caustic in real space, since they form a complex conjugate pair of roots to (144). The equal-phase paths of any pair of real rays do not intersect real space. However, the equal-phase surfaces for one real and one complex ray do intersect real space, and are shown in Fig. 31.

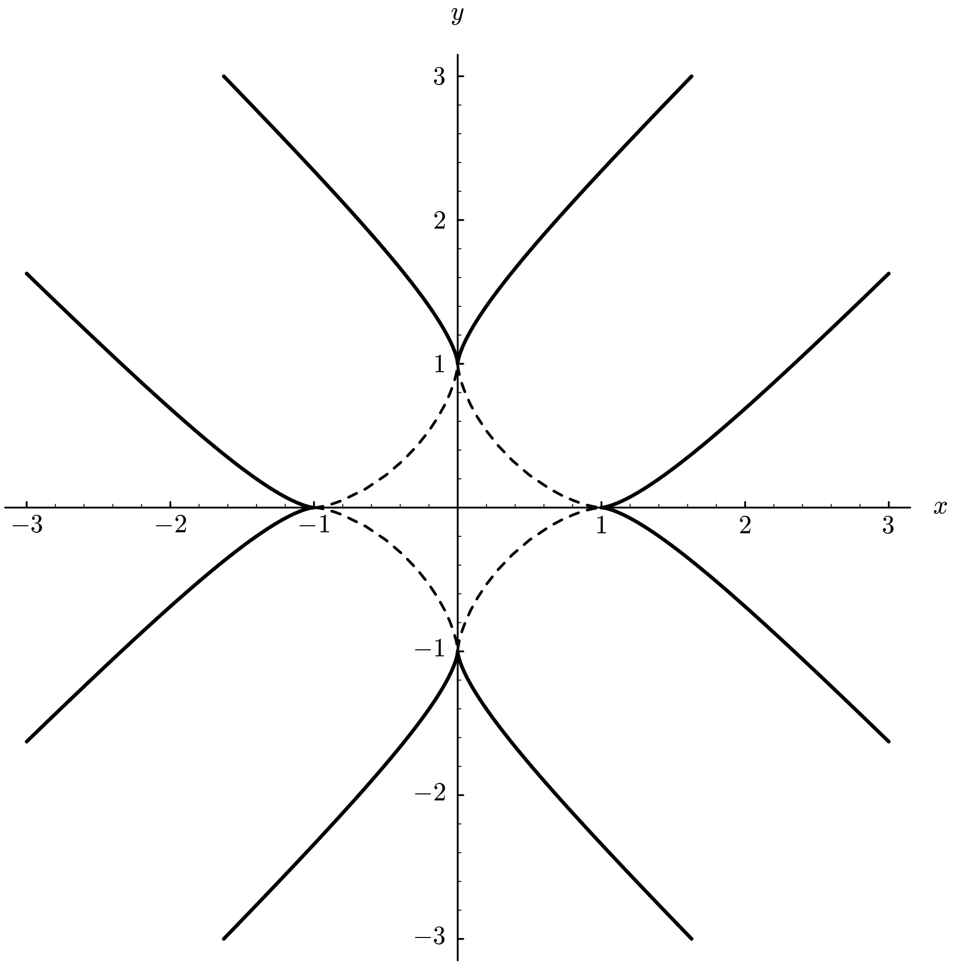


Figure 31: Equal-phase surfaces (solid curves) in problem (138)-(140). The caustic is also shown (dashed curve).

As we would expect, the real ray which has the same phase as the complex rays is ray number 4. The evidence so far strongly suggests that the portions of these last equal-phase surfaces emerging from $x = \pm 1$, $y = 0$ are indeed Stokes surfaces, and that we have the situation shown in Fig. 32.

To confirm this let us examine the Fourier transform solution of the problem (138)-(140). We find

$$\phi = -\sqrt{\frac{i}{2\pi k}} \int \exp(-i\lambda^2/(2k) - \sqrt{\lambda^2 - k^2}\, y + i\lambda x)\, d\lambda, \tag{145}$$

where we will identify the contour of integration and the branch of the square root shortly. We are interested in the asymptotic approximation to (145) as $k \to \infty$. We remove the branch point by making the substitution $\lambda = k\cos\theta$. Then

$$\phi = \sqrt{\frac{ik}{2\pi}} \int \sin\theta \exp(-ik[(\cos^2\theta)/2 + y\sin\theta - x\cos\theta])\, d\theta, \tag{146}$$

where the contour runs between two "valleys" at $n\pi \pm i\infty$, and will be specified shortly. There are no singularity or end point contributions, and so the saddle points correspond to the rays, as shown in §2.1; as expected there are four in the period $-\pi < \theta_1 < \pi$, where $\theta = \theta_1 + i\theta_2$, θ_1, θ_2 real. Now we can identify which contour we require by identifying the saddles with the rays, since we know which rays we expect to be present for x and y small. For such x and y the saddles are given by $\theta \sim y$, $\pi/2 - x$, $\pi - y$, $-\pi/2 + x$, giving ray contributions

$$\begin{aligned} u_1 &\sim -1/2 + x - y^2/2, &(147)\\ u_2 &\sim -y + x^2/2, &(148)\\ u_3 &\sim -1/2 - x - y^2/2, &(149)\\ u_4 &\sim y + x^2/2. &(150) \end{aligned}$$

Thus we require a contour which passes only through the saddle near $\theta = -\pi/2 + x$ for small x and y. This identifies the contour as that shown in Fig. 34(A).

Now let us vary x and y so as to circumnavigate the point $x = 1$, $y = 0$, and thus observe the switching on of the complex ray. We find the sequence of steepest descent paths shown in Figs. 34, 35 with the implications in real space indicated in Fig. 33; it is indeed only those components of the equal phase surface of Fig. 31 with cusps at $(1, 0)$ and $(-1, 0)$ which are active.

Notice the "double" Stokes surface in the analytic continuation of the solution into the lower half plane, on which the real ray turns on the second complex ray while itself being turned off by the first complex ray. The two complex saddles then coalesce at the lower caustic to form the one real outgoing ray. Note also that the solution in the vicinity of the lower caustic is different to that in the vicinity of the caustic in §3. Here, as in the "outgoing-only" caustic in §5, there are two complex rays on one side and only one real ray on the other.

Let us now see if we could deduce that the equal-phase contour of Fig. 32 is "active" if we had not had the Fourier transform solution (145) available.

We will consider the problem locally near $x = 1$, since if the Stokes line is active there, it will be active along its length by a continuity argument. Expanding

$$\begin{aligned} s &= 1 + \epsilon\bar{s}, &(151)\\ t &= \epsilon\bar{t}, &(152) \end{aligned}$$

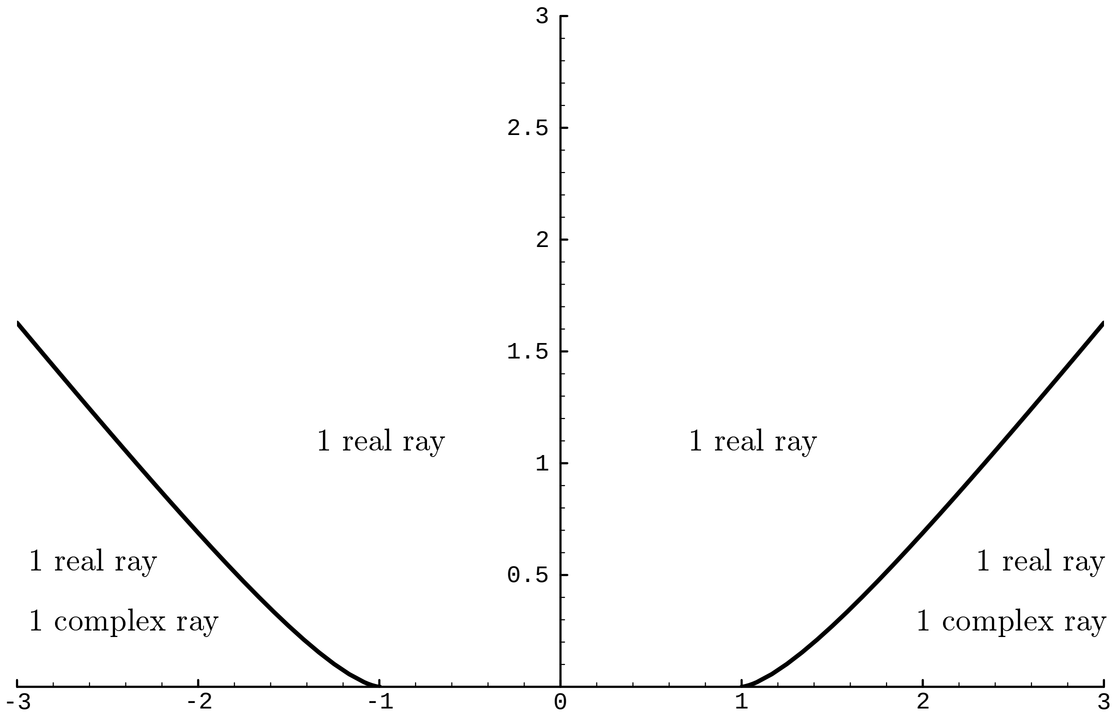


Figure 32: Rays contained in the solution of problem (138)-(140)

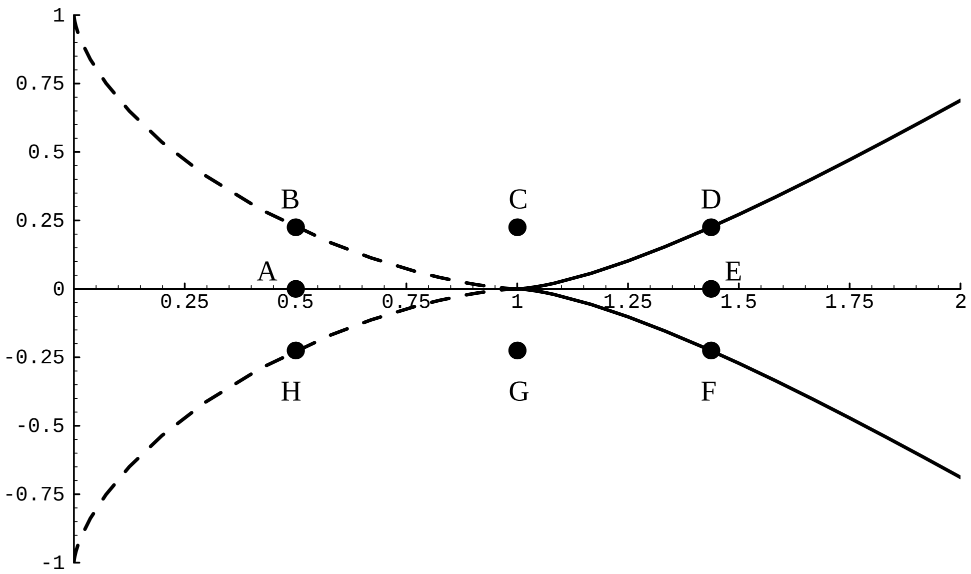
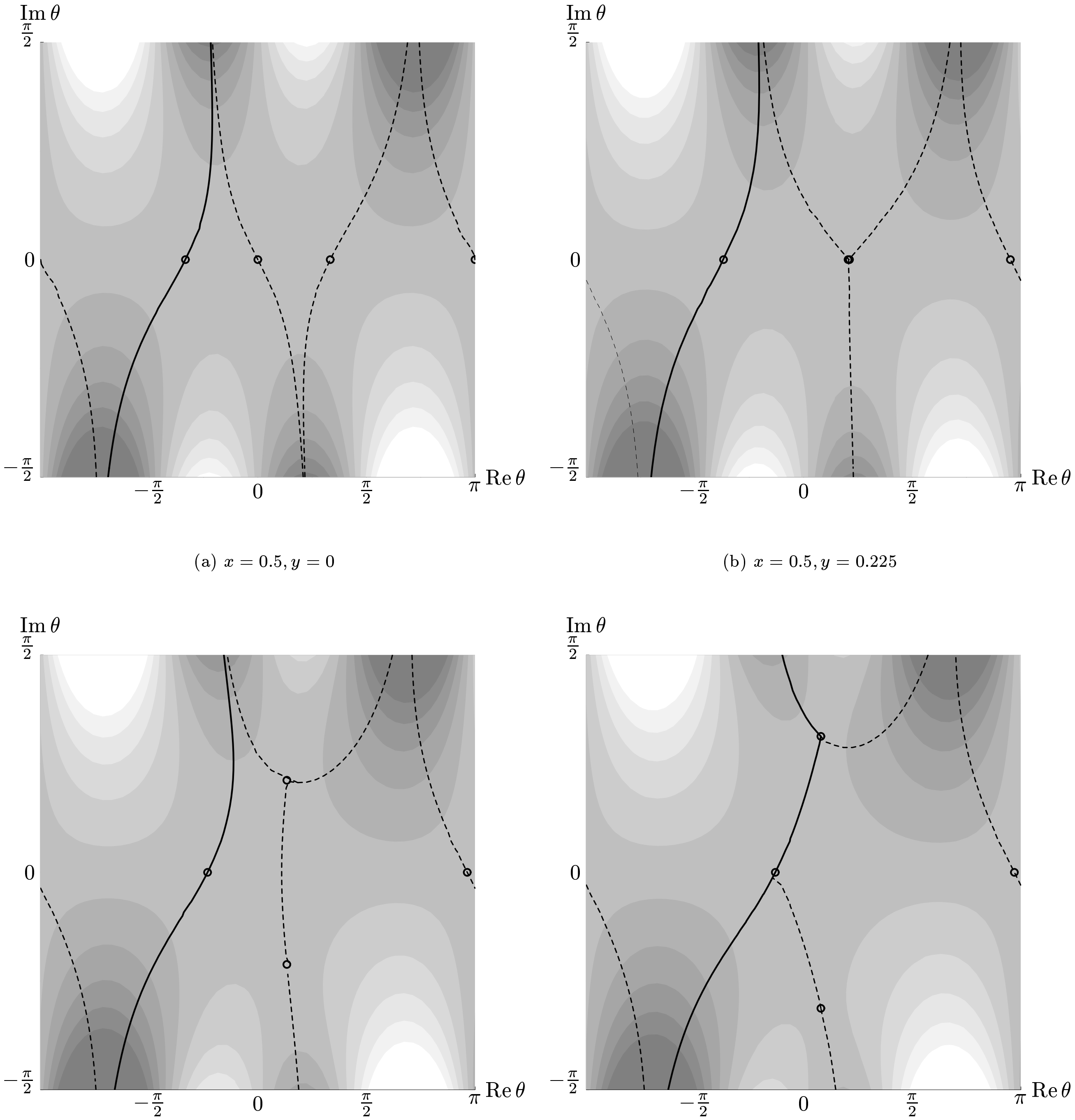


Figure 33: Diagram showing the points in real space corresponding to the steepest descent paths in Figs. 34 and 35. The caustic (dotted curve) and Stokes surface (solid curve) are shown.



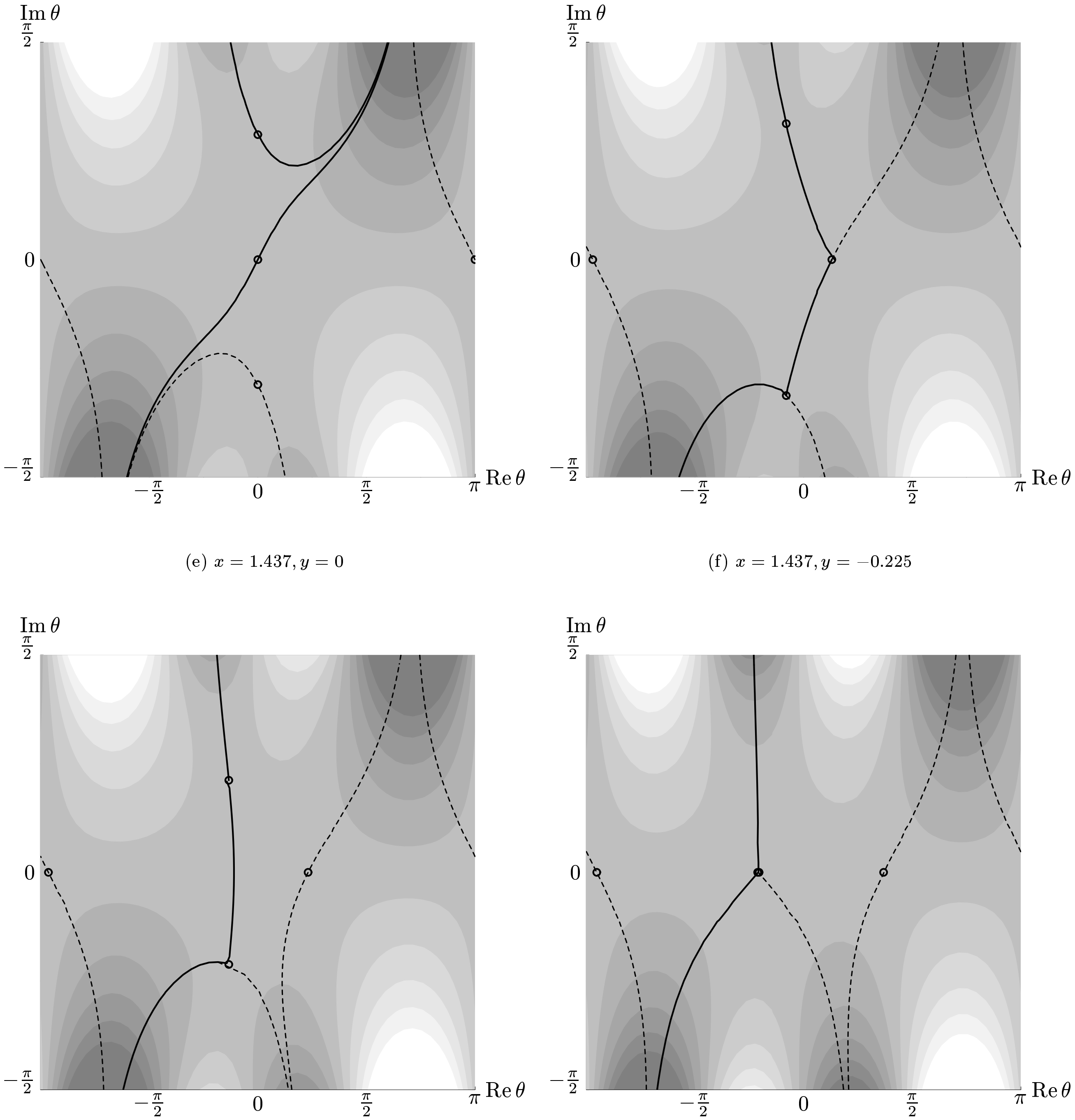
(a) $x = 0.5, y = 0$

(b) $x = 0.5, y = 0.225$

(c) $x = 1, y = 0.225$

(d) $x = 1.437, y = 0.225$

Figure 34: Steepest-descent paths for the model problem (138)–(140). Shading indicates "hills" (light) and "valleys" (dark) in the integrand; saddles are circles. The solid line is the integration path required for the solution to our problem; steepest-descent paths through other saddles are dotted. Saddles corresponding to real rays are on the line $\operatorname{Im}\theta = 0$.



(e) $x = 1.437, y = 0$

(f) $x = 1.437, y = -0.225$

(g) $x = 1, y = -0.225$

(h) $x = 0.5, y = -0.225$

Figure 35: Steepest-descent paths for the model problem (138)–(140). Shading indicates “hills” (light) and “valleys” (dark) in the integrand; saddles are circles. The solid line is the integration path required for the solution to our problem; steepest-descent paths through other saddles are dotted. Saddles corresponding to real rays are on the line $\operatorname{Im}\theta = 0$.

$$
\begin{aligned}
x &= 1 + \epsilon\bar{x}, & (153)\\
y &= \epsilon\bar{y}, & (154)\\
u - x &= \epsilon^2\bar{u}, & (155)
\end{aligned}
$$

gives the local ray solution as

$$
\begin{aligned}
\bar{x} &= \bar{s} + \bar{t}, & (156)\\
\bar{y} &= \sqrt{2}\,\bar{t}(-\bar{s})^{1/2}, & (157)\\
\bar{u} &= \bar{s}(\bar{s} - 2\bar{t})/2. & (158)
\end{aligned}
$$

There are now three participating rays; ray number 3 in Fig. 30 plays no part. Let us write $\eta = \bar{t}/\bar{s}$, for reasons which will become clear later. Then

$$
\begin{aligned}
\bar{x} &= \bar{s}(1 + \eta), & (159)\\
\bar{y}^2 &= -2\bar{s}^3\eta^2, & (160)\\
\bar{u} &= \frac{\bar{s}^2(1 - 2\eta)}{2} = \frac{\bar{x}^2(1 - 2\eta)}{2(1 + \eta)^2}. & (161)
\end{aligned}
$$

Eliminating $\bar{s}$ gives

$$(1 + \eta)^3 + 2\eta^2(\bar{x}^3/\bar{y}^2) = 0. \qquad (162)$$

Thus, for real $\bar{x}$ and $\bar{y}$, the three rays are given by the three roots of equation (162), and the equal-phase surfaces are given by

$$\mathrm{Re}\left(\frac{(1 - 2\eta_1)}{2(1 + \eta_1)^2}\right) = \mathrm{Re}\left(\frac{(1 - 2\eta_2)}{2(1 + \eta_2)^2}\right) \qquad (163)$$

where η_1 and η_2 are any two roots. Let us denote $\bar{x}^3/\bar{y}^2$ by a. We wish to find the real values of a for which (163) holds. Denoting the real root of (162) by η_1, we find the other roots can be written as

$$
\begin{aligned}
\eta_2 &= \frac{3\eta_1 + 1 + \sqrt{(3\eta_1)^2 + 4\eta_1^3}}{2\eta_1^2}, & (164)\\
\eta_3 &= \frac{3\eta_1 + 1 - \sqrt{(3\eta_1)^2 + 4\eta_1^3}}{2\eta_1^2}. & (165)
\end{aligned}
$$

The equal-phase contour we require is given by

$$\mathrm{Re}\left(\frac{(1 - 2\eta_1)}{2(1 + \eta_1)^2} - \frac{(1 - 2\eta_2)}{2(1 + \eta_2)^2}\right) = \frac{1 - 10\eta_1 - 2\eta_1^2}{2(1 + \eta_1)^2} = 0, \qquad (166)$$

when η_1 is real and less than $-1/4$. Thus the equal-phase surfaces correspond to $\eta_1 = -5/2 - 3\sqrt{3}/2$, giving $a = 27/(10 + 6\sqrt{3})$; hence they are locally the semicubical parabolae

$$\bar{x}^3 = \frac{27\bar{y}^2}{2(5 + 3\sqrt{3})}. \qquad (167)$$

Let us now examine the behaviour of the amplitude coefficients A_n. We expect that

$$A_n \sim \frac{\Lambda\Gamma(n + \gamma)}{f^n} \qquad \text{as } n \to \infty. \qquad (168)$$

Substituting this expression into equation (25) and retaining only the leading order terms in n as $n \to \infty$ gives, in the notation of (24)

$$-2f_\tau = \alpha^2 f_\sigma^2 + f_\tau^2 + \frac{2\alpha^2(\beta' - u_0')f_\sigma f_\tau}{\tau^2} + \frac{\alpha^2(\beta' - u_0')^2 f_\tau^2}{\tau^2}. \tag{169}$$

Now let us look locally near $\sigma = 1$, $\tau = 0$ by rescaling

$$\begin{aligned} \sigma &= 1 + \epsilon\bar{\sigma}, && (170)\\ \tau &= \epsilon\bar{\tau}, && (171)\\ f &= \epsilon^2 \bar{f}. && (172) \end{aligned}$$

We find that to leading order in ϵ

$$\bar{f}_{\bar{\tau}} = \frac{\bar{\sigma}}{\bar{\tau}^2}\left(\bar{f}_{\bar{\sigma}} - 3\bar{f}_{\bar{\tau}}\right)^2, \tag{173}$$

or, since $\bar{s} = \bar{\sigma}$, $\bar{t} = \bar{\tau} + 2\bar{\sigma}$

$$\bar{f}_{\bar{t}} = \frac{\bar{s}}{(\bar{t} - 2\bar{s})^2}\left(\bar{f}_{\bar{s}} - \bar{f}_{\bar{t}}\right)^2. \tag{174}$$

The scaling (170)-(172) suggests that we seek a similarity solution of the form $\bar{f} = \bar{s}^2 F(\eta)$, where $\eta = \bar{t}/\bar{s}$ as before, giving

$$(2-\eta)^2 F' = \left(2F - (\eta+1)F'\right)^2, \tag{175}$$

where $'$ represents $d/d\eta$. We know that the active Stokes surface will be given by i/f real and positive. Thus we are interested in the surfaces $\mathrm{Re}(f) = 0$, i.e. $\mathrm{Re}(\bar{s}^2 F) = 0$. Since $\bar{s} = \bar{x}/(1+\eta)$, and we are interested in real $\bar{x}$, we are therefore seeking the surfaces $\mathrm{Re}(F/(1+\eta)^2) = 0$. This motivates the substitution $F = (1+\eta)^2 g$, giving

$$(1+\eta)^5 (g')^2 - (2-\eta)^2(1+\eta)g' - 2(2-\eta)^2 g = 0. \tag{176}$$

The boundary condition on (176) comes from consideration of the exact solution for A_n, namely the series solution (30). Since the coefficient $g_{n,n,3n}$ depends only on $g_{n-1,n-1,3n-3}$ we can solve for it exactly. Letting $b_n = g_{n,n,3n}$ for ease of notation we have

$$\begin{aligned} b_{n+1} &= -\frac{3s^2(1-s^2)(3n+5/2)(3n+1/2)b_n}{2(n+1)}, \\ &\sim \frac{3\epsilon\bar{s}(3n+5/2)(3n+1/2)b_n}{(n+1)} \end{aligned}$$

as $\epsilon \to 0$. Hence $b_n \sim B(n,s)3^{3n}\epsilon^n \bar{s}^n n!$ as $n \to \infty$, where $|\log(B)| = o(n)$. Now, as $\tau \to 0$, $g_{n,n,3n}$ will dominate the behaviour of A_n, giving

$$\begin{aligned} A_n &\sim B(n,s)\frac{3^{3n}\bar{s}^n n!}{\epsilon^{2n}\bar{\tau}^{3n}}, \\ &\sim \frac{B(n,s)}{\epsilon^{2n}\bar{s}^{2n}}\left(\frac{3}{(\eta-2)}\right)^{3n} n! \end{aligned}$$

as $n \to \infty$, $\bar{\tau} \to 0$. Hence we have the boundary condition

$$\bar{f} \sim \frac{(\eta-2)^3}{27} \qquad \text{as } \eta \to 2, \tag{177}$$

giving

$$g \sim \frac{(\eta-2)^3}{3^5} \qquad \text{as } \eta \to 2. \tag{178}$$

The solution of (176) satisfying (178) is

$$g = -\frac{1}{2}\left[\frac{1-2\eta}{(1+\eta)^2} - \frac{4\eta^2(-1-3\eta+\eta^2-(1+\eta)\sqrt{1+4\eta})}{(1+3\eta+2\eta^2+(1+\eta)\sqrt{1+4\eta})^2}\right]. \tag{179}$$

The contour $\operatorname{Re} g = 0$ is given by $1 - 10\eta - 2\eta^2 = 0$, for η real and less than $-1/4$, giving $\eta = -5/2 - 3\sqrt{3}/2$ as expected.

6.2 Total Internal Reflection

Let us now examine the problem of total internal reflection which we introduced in §2.3, but with a line source rather than plane wave incidence. We have

$$\begin{aligned}
\nabla^2\phi_1 + k_1^2\phi_1 &= \delta(x)\delta(y-h) \qquad y > 0, && (180)\\
\nabla^2\phi_2 + k_2^2\phi_2 &= 0 \qquad y < 0, && (181)\\
\phi_1 &= \phi_2 \qquad \text{on } y = 0, && (182)\\
\frac{\partial \phi_1}{\partial y} &= \lambda\frac{\partial \phi_2}{\partial y} \qquad \text{on } y = 0. && (183)
\end{aligned}$$

A line source at the point $(0,h)$, $h>0$, gives an incoming wavefield

$$\begin{aligned}
\phi_{\mathrm{inc}} &= H_0^{(1)}(k_1(x^2+(y-h)^2)^{1/2})\\
&\sim \frac{1}{\sqrt{k_1}(x^2+(y-h)^2)^{1/4}}e^{ik_1(x^2+(y-h)^2)^{1/2}}.
\end{aligned}$$

Ray Solution Let us first examine the ray solution of the problem. Writing

$$\begin{aligned}
\phi_1 - \phi_{\mathrm{inc}} &= A^{(1)}e^{ik_1u_1}, && (184)\\
\phi_2 &= A^{(2)}e^{ik_2u_2}, && (185)
\end{aligned}$$

we find that in order to satisfy the continuity conditions at the interface we must have

$$k_1u_1 = k_2u_2 = k_1(x^2+h^2)^{1/2}. \tag{186}$$

Hence we have the ray solution

$$\begin{aligned}
x &= s + \frac{st}{(s^2+h^2)^{1/2}}, && (187)\\
y &= \frac{ht}{(s^2+h^2)^{1/2}}, && (188)\\
u_1 &= (s^2+h^2)^{1/2} + t, && (189)
\end{aligned}$$

in $y > 0$, and

$$\begin{aligned}
x &= s + \frac{nst}{(s^2+h^2)^{1/2}}, && (190)\\
y &= -\frac{(h^2-s^2(n^2-1))^{1/2}t}{(s^2+h^2)^{1/2}}, && (191)\\
u_2 &= n(s^2+h^2)^{1/2} + t, && (192)
\end{aligned}$$

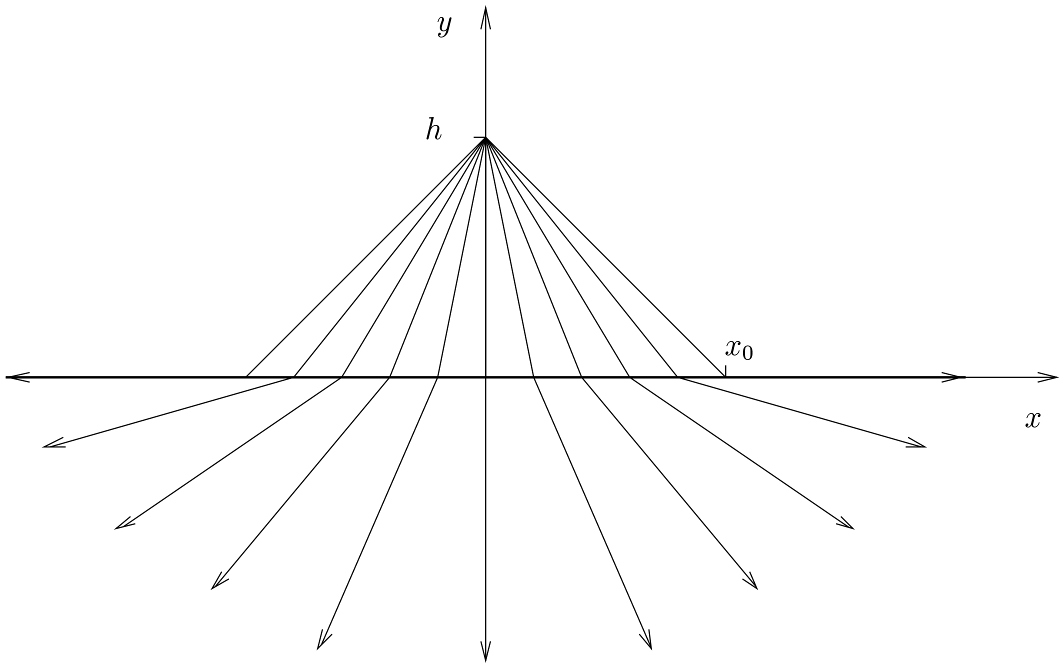


Figure 36: Line source incidence in the problem of total internal reflection (180)-(183). Only the incident and transmitted rays are shown for clarity.

in $y < 0$, where $n = k_1/k_2 > 1$. Let us consider the solution in the lower medium, and its analytic continuation into $y > 0$. Eliminating t we find

$$(x-s)^2(h^2 - s^2(n^2-1)) = n^2y^2s^2. \tag{193}$$

This is a quartic equation for s, and hence there will be four rays through any one point. To identify these four rays let us find the caustics associated with them. We find that β is given by

$$\beta = \frac{(s^2+h^2)^{1/2}(h^2 - s^2(n^2-1))}{h^2 n}, \tag{194}$$

so that the equation of the caustic $t + \beta = 0$ is given parametrically by

$$x = \frac{s^3(n^2-1))}{h^2}, \tag{195}$$

$$y = \frac{(h^2 - s^2(n^2-1))^{3/2}}{h^2 n}, \tag{196}$$

which is shown in Fig. 37.

As in the previous problem, we can identify the rays as the tangents to this curve. Inside the caustic all the rays are real, while outside the caustic there are two real and two complex rays.

We see that the only ray which should be present inside the caustic in the physical solution is ray number 2. However, as before, this ray will not satisfy the boundary data for $|x| > x_0 = h/\sqrt{n^2-1}$, $y = 0$, and so, close to $x > x_0$, $y = 0$, one of the complex rays 1 and 4 must be present (the smaller one). Thus there must be a Stokes surface across which a complex ray is switched on/off somewhere outside the lower caustic (since the complex ray is not present at the caustic). If we plot the equal-phase surfaces of the rays we again find

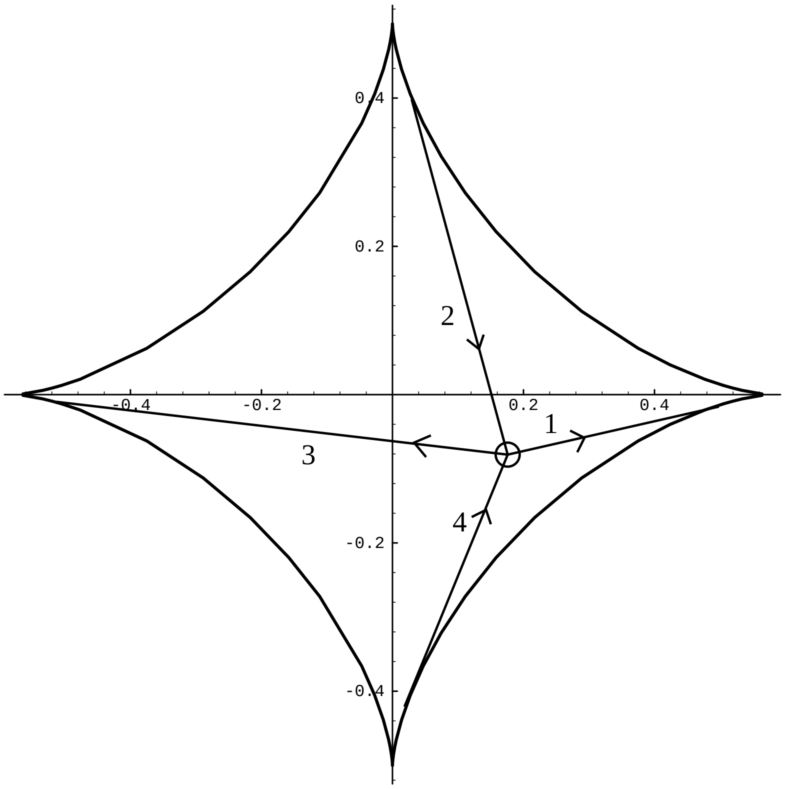


Figure 37: Caustic in the problem of total internal reflection (180)-(183) with $n = 2$, $h = 1$. The caustic in the upper halfspace is that which is generated by analytically continuing the transmitted ray back into this halfspace. The caustic in the lower halfspace is the reflection of this. Neither will be present in the physical solution.

that the complex rays, coming from complex conjugate roots of (193), have equal phase wherever they exist. Again, for any pair of real rays the equal-phase surfaces do not intersect real space but for the case of one real ray and one complex ray we find the equal-phase surfaces shown in Fig. 38. Thus we expect that the complex ray is switched off across the curves shown in Fig. 39, exactly as in the previous problem.

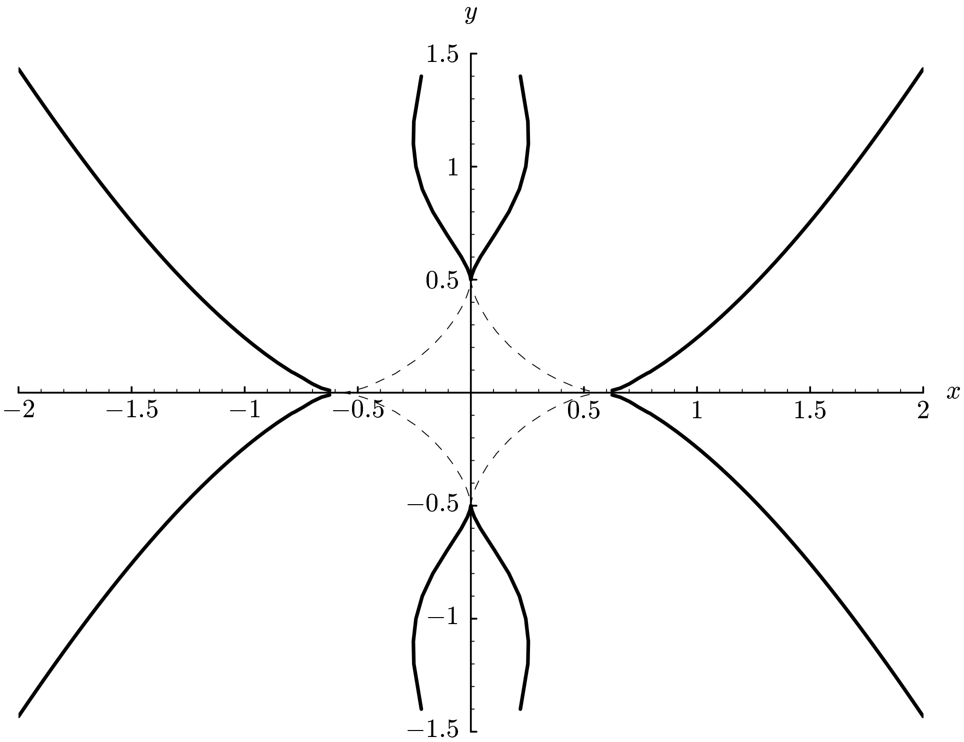


Figure 38: Equal phase surfaces (solid curves) for the problem (180)-(183). The caustic (dashed) is also shown.

Let us again confirm this by considering the Fourier transform solution. Let $\phi_{\text{ref}} = \phi_1 - \phi_{\text{inc}}$. Then we have

$$\begin{aligned} \bar{\phi}_{\text{inc}} &= \frac{e^{il_1|y-h|}}{2il_1}, && (197) \\ \bar{\phi}_{\text{ref}} &= Re^{il_1(h+y)}, && (198) \\ \bar{\phi}_2 &= Te^{il_2(h-y)}, && (199) \end{aligned}$$

where $l_i = \sqrt{k_i^2 - w^2}$, $i = 1, 2$, and

$$\begin{aligned} R &= \frac{1}{2il_1}\frac{\lambda l_2 - l_1}{\lambda l_2 + l_1}, && (200) \\ T &= -\frac{i}{\lambda l_2 + l_1}. && (201) \end{aligned}$$

We start by performing a steepest descents analysis on the inversion integral for ϕ_2. Let $w = k_1\bar{w}$. Then

$$\phi_2 = \int g(\bar{w})e^{ik_1 f(\bar{w})}\, d\bar{w} \tag{202}$$

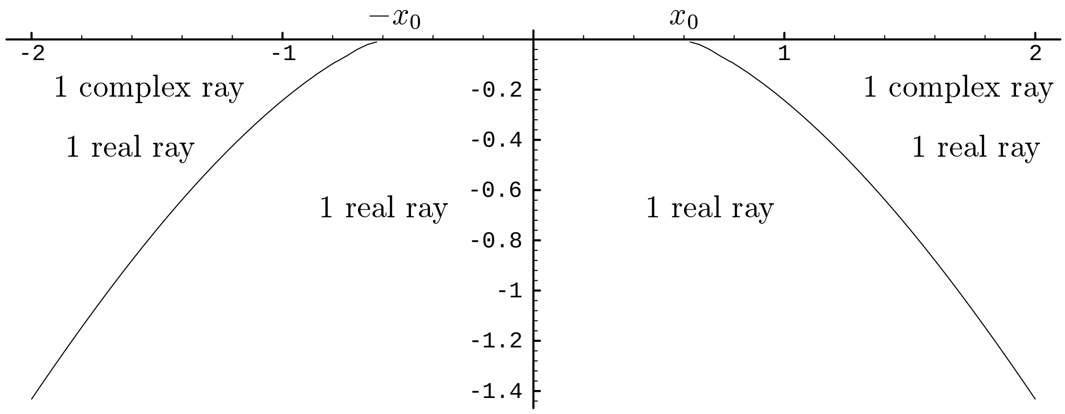


Figure 39: Rays contained in the solution of the problem (180)-(183).

where the contour is to be determined, and

$$\begin{aligned} f(\bar{w}) &= -y\sqrt{1/n^2 - \bar{w}^2} + x\bar{w} + h\sqrt{1-\bar{w}^2}, & (203)\\ g(\bar{w}) &= \frac{-i}{2\pi\left(\lambda\sqrt{1/n^2-\bar{w}^2}+\sqrt{1-\bar{w}^2}\right)}. & (204)\end{aligned}$$

Let us make the change of variable $s = \bar{w}/\sqrt{1-\bar{w}^2}$. This is motivated by the ray solution, and moves the branch points at ± 1 to infinity, while infinity moves to the branch points $\pm i$. Thus we expect our contour now to start and finish at these branch points. We find

$$f(\bar{w}) = \tilde{f}(s) = \frac{1}{\sqrt{1+s^2}}\left(h - \frac{y}{n}(1-(n^2-1)s^2)^{1/2} + sx\right). \tag{205}$$

Let us now remove the branch point at $s = 1/\sqrt{n^2-1}$ by setting $s = (1/\sqrt{n^2-1})\cos\theta$. We find

$$\tilde{f}(s) = \hat{f}(\theta) = \left(\frac{n^2-1}{n^2-\sin^2\theta}\right)^{1/2}\left(h - \frac{y}{n}\sin\theta + \frac{x\cos\theta}{\sqrt{n^2-1}}\right). \tag{206}$$

We can now clearly see the interaction of the saddles in the θ-plane. The steepest descent contours are plotted in Figs. 41,42 as x and y vary about the point $(x_0, 0)$, and the implications in physical space are plotted in Fig. 40. The behaviour of the steepest descent contour is qualitatively the same as in the previous example, as might have been expected. We see that the complex saddle is switched on across the Stokes surface as expected.

We find that g is given by

$$g(\bar{w})d\bar{w} = \hat{g}(\theta)d\theta = -\frac{1}{2\pi i}\frac{(n^2-1)^{1/2}\sin\theta}{n^2-\sin^2\theta}\left(\frac{1}{1+(\lambda/n)\sin\theta}\right)d\theta. \tag{207}$$

Note that the branch point has been removed from g also, and that we are left only with a pole. Since this pole lies on the negative real axis, we see from Figs. 41, 42 that the steepest descent path never crosses it for real x and y, so that its corresponding rayfield never contributes in real space.

Let us now examine the reflected wavefield. Here it is well-known that the interaction between the steepest descent contour and the branch cut in $g(\bar{w})$ leads to the switching on of a head wave [7, 41]. If we perform the same substitutions on ϕ_{ref} as we did on ϕ_2 we again remove the branch cut from g, and the head wave arises more conveniently as a saddle contribution. We have

$$\phi_{\text{ref}} = \int_{-\infty}^{\infty} g(\bar{w})e^{ik_1 f(\bar{w})}\,d\bar{w}, \tag{208}$$

with

$$\begin{aligned} f(\bar{w}) &= (y+h)\sqrt{1-\bar{w}^2} + x\bar{w}, & (209)\\ g(\bar{w})d\bar{w} &= \frac{1}{4\pi i}\frac{1}{\sqrt{1-\bar{w}^2}}\left(\frac{-\sqrt{1-\bar{w}^2}+\lambda\sqrt{1/n^2-\bar{w}^2}}{\sqrt{1-\bar{w}^2}+\lambda\sqrt{1/n^2-\bar{w}^2}}\right)d\bar{w}. & (210)\end{aligned}$$

Using the same transformations as before we find

$$\phi_{\text{ref}} = \int \hat{g}(\theta)e^{ik_1\hat{f}(\theta)}\,d\theta, \tag{211}$$

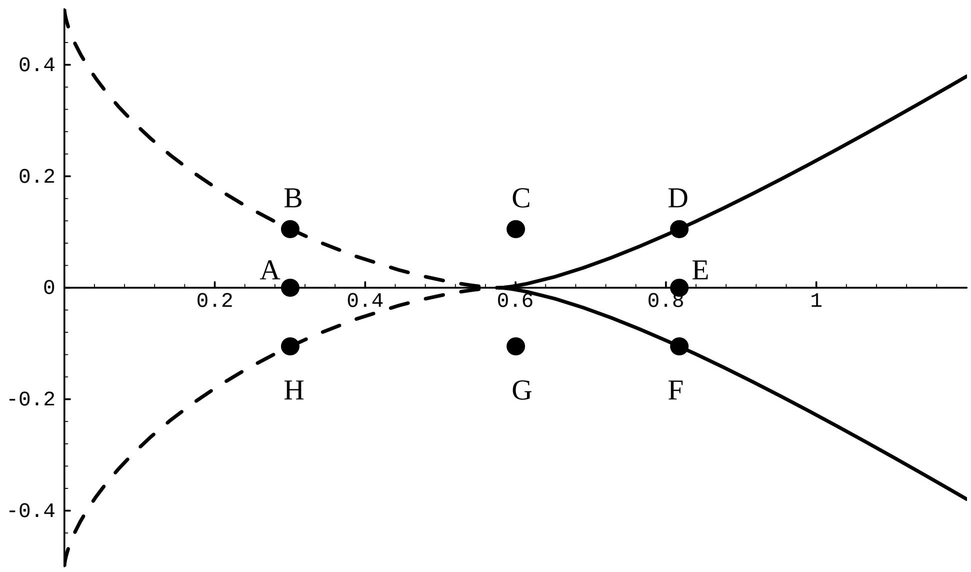
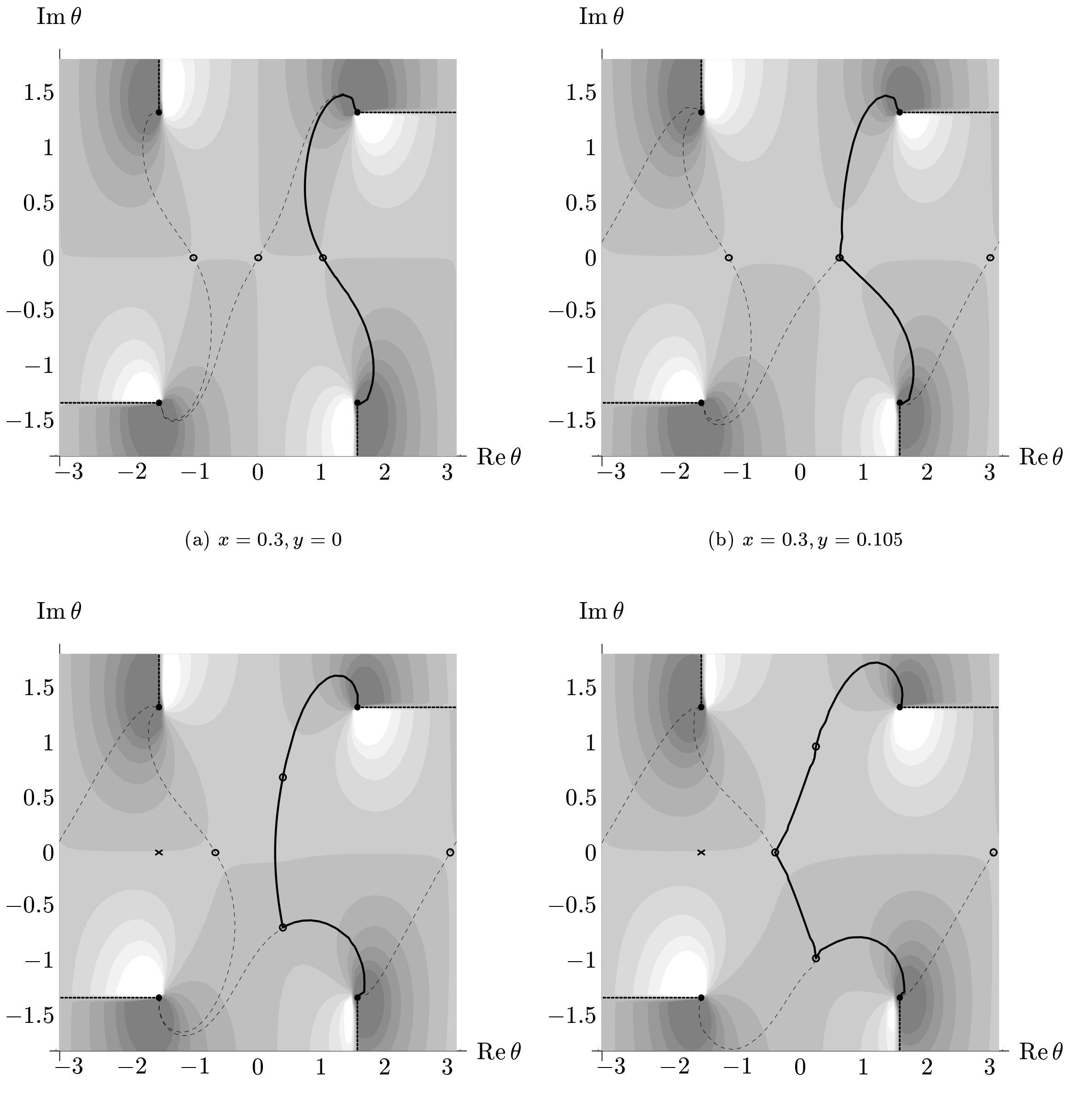


Figure 40: Diagram showing the points in space corresponding to the steepest descent paths in Figs. 41 and 42. The dotted line indicates the caustic, and the solid line indicates the Stokes surface. Note that the upper halfspace represents the analytic continuation of the solution in the lower medium, and not the solution in the upper medium.



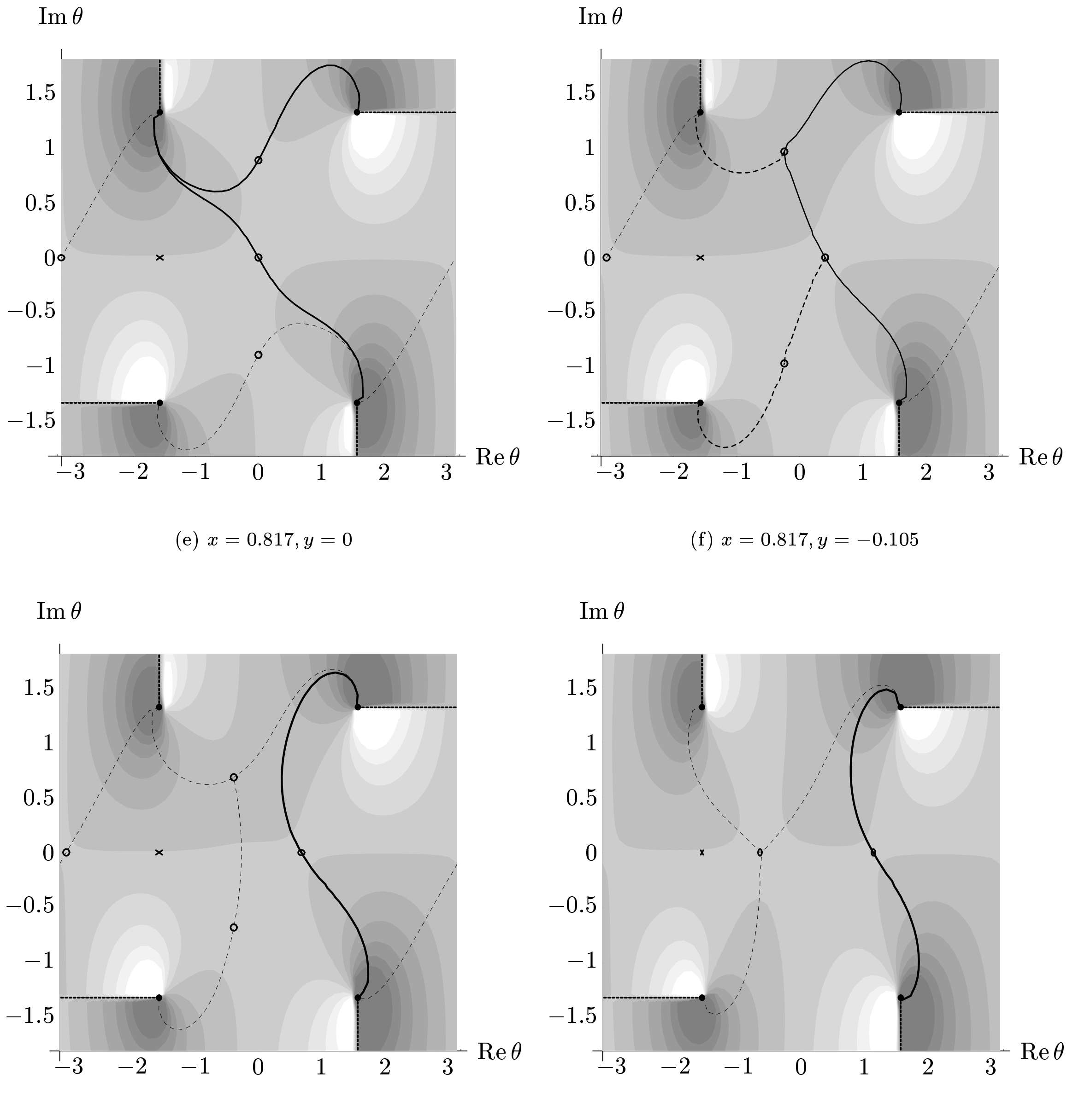
(a) $x = 0.3, y = 0$

(b) $x = 0.3, y = 0.105$

(c) $x = 0.6, y = 0.105$

(d) $x = 0.817, y = 0.105$

Figure 41: Steepest-descent paths for ϕ_2 in the problem (180)–(183). Shading indicates "hills" (light) and "valleys" (dark) in the integrand; saddles are circles. The solid line is the integration path required for the solution to our problem; steepest-descent paths through other saddles are dotted. Saddles corresponding to real rays are on the line $\operatorname{Im}\theta = 0$.



(e) $x = 0.817, y = 0$

(f) $x = 0.817, y = -0.105$

(g) $x = 0.6, y = -0.105$

(h) $x = 0.3, y = -0.105$

Figure 42: Steepest-descent paths for ϕ_2 in the problem (180)–(183). Shading indicates "hills" (light) and "valleys" (dark) in the integrand; saddles are circles. The solid line is the integration path required for the solution to our problem; steepest-descent paths through other saddles are dotted. Saddles corresponding to real rays are on the line $\mathrm{Im}\,\theta = 0$.

where

$$\hat{f}(\theta) = \frac{(n^2-1)^{1/2}(y+h)+x\cos\theta}{(n^2-\sin^2\theta)^{1/2}}, \tag{212}$$

$$\hat{g}(\theta)d\theta = \frac{1}{4\pi i}\frac{(n^2-1)^{1/2}\sin\theta}{n^2-\sin^2\theta}\frac{(\lambda\sin\theta-n)}{(\lambda\sin\theta+n)}\,d\theta, \tag{213}$$

where the contour of integration must now start and finish at the roots of $\cos\theta = \pm i\sqrt{n^2-1}$ as before; the exact contour will be determined shortly. There is now no branch point in g, but there are four saddles in f. Solving $f'(\theta)=0$ we find

$$\sin\theta = 0 \qquad \text{or} \qquad \cos\theta = \frac{x\sqrt{n^2-1}}{y+h}.$$

The second of these gives the reflected field as required. The first gives the contribution

$$\phi \sim Ae^{ik_1\left[(n^2-1)^{1/2}(y+h)/n+x/n\right]}, \tag{214}$$

where

$$A = \frac{\lambda e^{3\pi i/4}n^{3/2}}{\sqrt{2\pi}k_1^{3/2}(n^2-1)^{1/4}(x\sqrt{n^2-1}-(y+h))^{3/2}} \tag{215}$$

and is the head wave. In Figs. 44, 45 we plot the steepest descent path as x and y vary around the point $(x_0, 0)$ (again we analytically continue the solution into $y<0$), with the resulting waves in real space shown in Fig. 43. Thus we see that the head wave will be switched on across a Stokes surface in complex space. In this case the intersection of the Stokes surface with real space is the point at which the saddles coalesce.

6.3 Plane wave scattering by a cylinder

Let us now consider the problem of the complexification of the shadow boundary of a smooth scatterer. We will examine first the scattering of a real plane wave by a cylinder $\Omega \subset \mathbb{R}^2$, for which the asymptotic solution in real space is well known [21, 19]. In the following section we will consider the scattering of a complex plane wave by a cylinder, for which the point of tangency is in complex space.

We have

$$\nabla^2\phi + k^2\phi = 0, \qquad (x,y)\in\mathbb{R}^2\backslash\Omega, \tag{216}$$

$$\phi = 0, \qquad (x,y)\in\partial\Omega, \tag{217}$$

with an incoming field $\phi = e^{ikx}$ as $x\to-\infty$ and a suitable radiation condition to be imposed on the scattered field. Parametrising the cylinder $x = x_0(s)$, $y = y_0(s)$ as usual gives the ray solution for the reflected field as

$$x = x_0(s) + \left((x_0'(s))^2 - (y_0'(s))^2\right)t, \tag{218}$$

$$y = y_0(s) + 2x_0'(s)y_0'(s)t, \tag{219}$$

$$u = x_0(s) + t. \tag{220}$$

We find that from these equations and (14),

$$\beta(s) = \frac{y_0'}{2(x_0''y_0' - y_0''x_0')}. \tag{221}$$

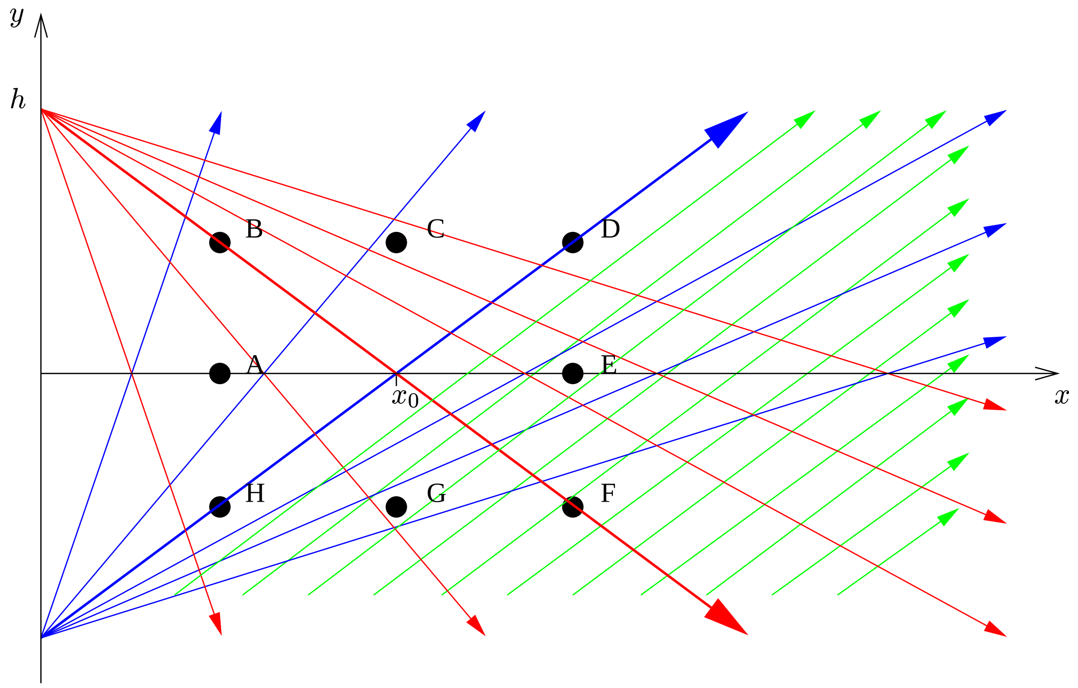
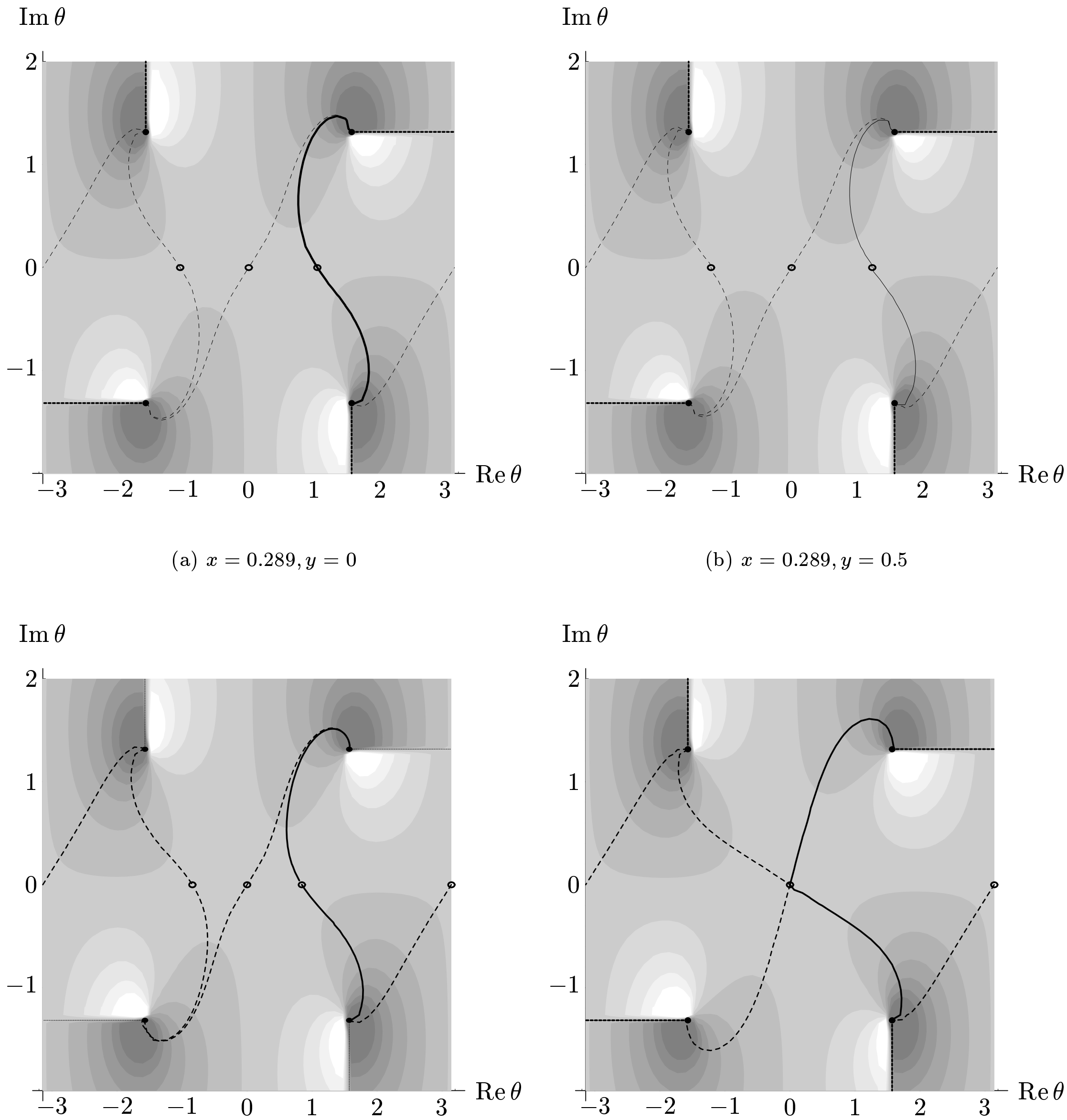


Figure 43: Diagram showing the points in space corresponding to the steepest descent paths in Figs. 44 and 45. The green, red and blue rays are the head wave, the incident rays and the reflected rays respectively. The bold ray is the critically incident ray across which the head wave is switched on. Note that the lower halfspace represents the analytic continuation of the solution in the upper medium, and not the solution in the lower medium.



(a) $x = 0.289, y = 0$

(b) $x = 0.289, y = 0.5$

(c) $x = 0.577, y = 0.5$

(d) $x = 0.866, y = 0.5$

Figure 44: Steepest-descent contours for ϕ_{ref} in the problem (180)-(183).

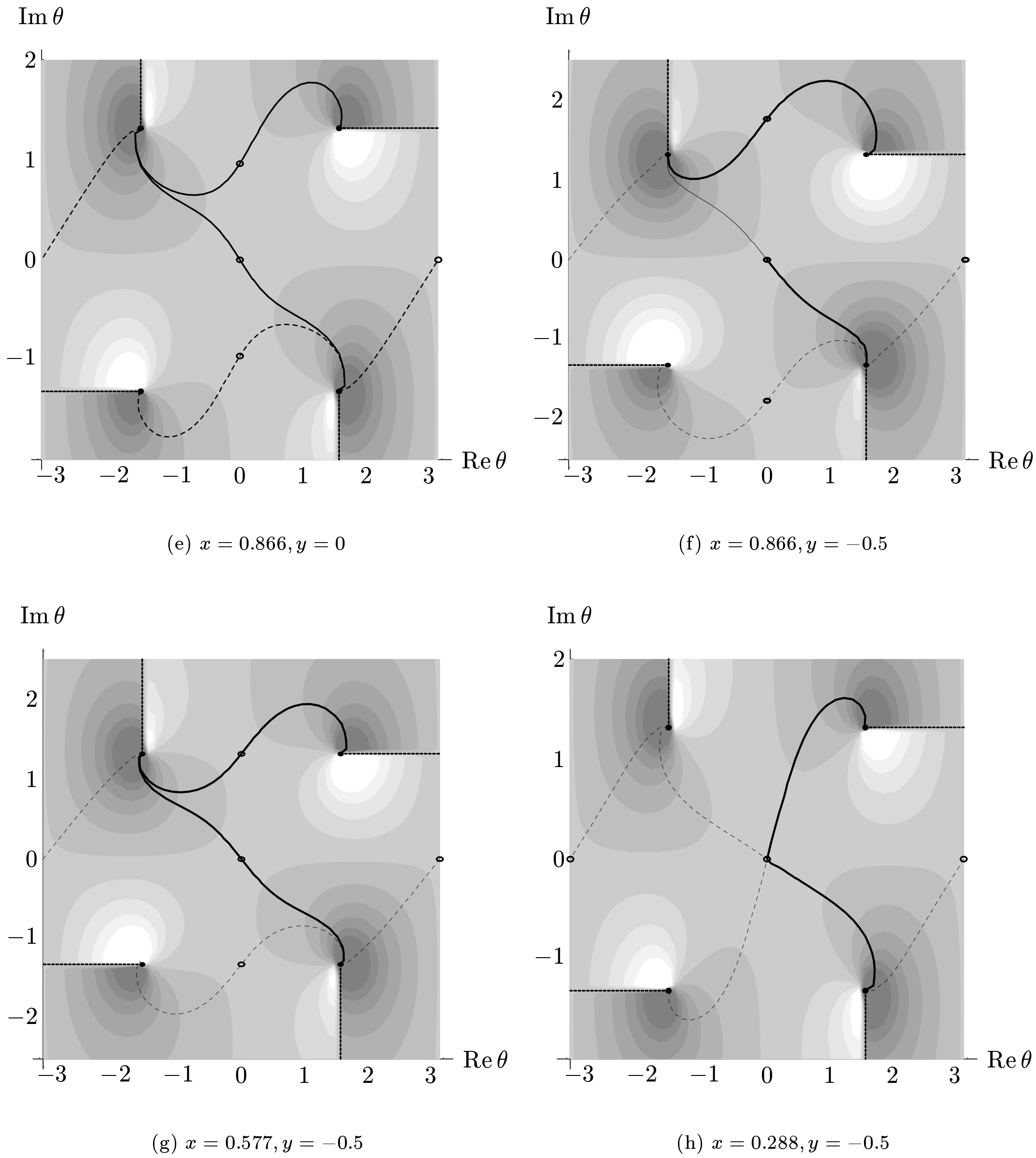


Figure 45: Steepest-descent contours for ϕ_{ref} in the problem (180)-(183).

At the top of the cylinder $y_0' = 0$, $x_0' \neq 0$, so that $\beta = 0$, corresponding to a tangent ray. Without loss of generality we take this point to be the origin $x = 0$, $y = 0$, and parametrise the boundary so that this point corresponds to $s = 0$.

In [21, 29] it has been proposed that wherever an incoming ray is externally tangent to a smooth convex boundary, it initiates a special type of surface ray, called a surface creeping ray, which propagates along the boundary of the scatterer into its unlit side. As this ray propagates around the body it continuously sheds further rays that leave the boundary tangentially, which we shall call creeping rays. With this geometrical prescription we can see that the boundary must be a caustic for the shed rays and hence that the creeping ray phenomenon can be thought of as a boundary layer effect. Corroboration of this idea in the form of uniform asymptotic theories has been provided in [2, 30]. In order to adopt a systematic asymptotic approach, it is necessary to perform a detailed local analysis of the wave structure close to the point of tangency. This demands the inner scalings

$$\hat{n} = \kappa_0^{1/3} k^{2/3} n, \qquad \hat{s} = \kappa_0^{2/3} k^{1/3} s, \tag{222}$$

for the normal and tangential coordinates n and s respectively, where κ_0 is the curvature of the body at the point of tangency, assumed to be of order one (see Figs. 46, 47).

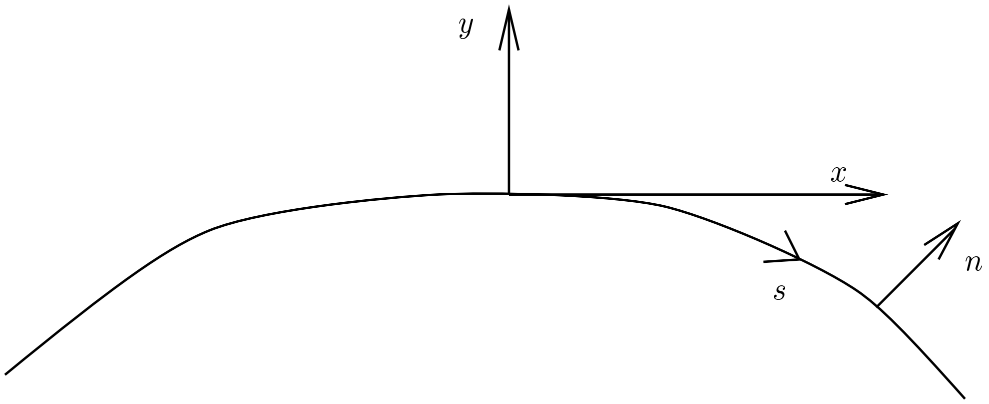


Figure 46: Local coordinate system near a point of tangency.

If we insert these scalings into the Helmholtz equation and seek a solution of the form

$$\phi = e^{iks}\hat{A}(\hat{s}, \hat{n}) + e^{ikx}, \tag{223}$$

motivated by the structure of the incoming field e^{ikx} close to the point of tangency, we find that the leading-order diffracted amplitude $\hat{A}$ satisfies the Fock-Leontovich equation [28, 45, 46]

$$\frac{\partial^2 \hat{A}}{\partial \hat{n}^2} + 2i\frac{\partial \hat{A}}{\partial \hat{s}} + 2\hat{n}\hat{A} = 0, \qquad \hat{n} > 0, \tag{224}$$

subject to the conditions

$$\hat{A}(\hat{s}, 0) = -e^{-i\hat{s}^3/6}, \tag{225}$$

and

$$e^{ik\hat{s}}\hat{A}(\hat{s}, \hat{n}) \text{ outgoing as } \hat{n} \to \infty. \tag{226}$$

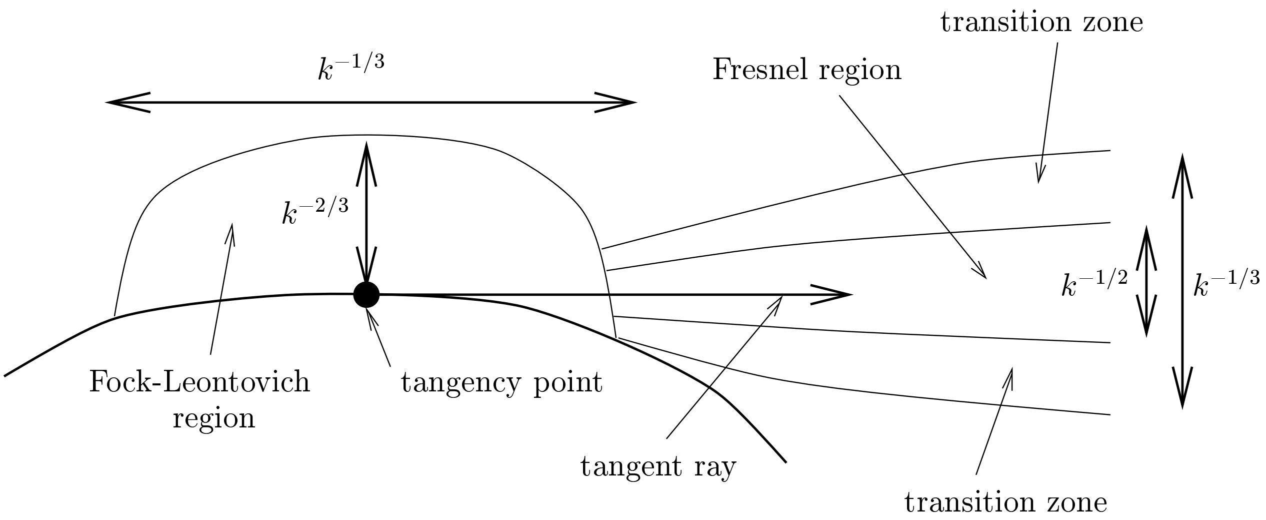


Figure 47: Diagram showing the tangency point, Fock-Leontovich region, shadow boundary and transition zones.

We can see immediately from (224) that if $n = O(k^{-2/3})$ but $s = O(1)$, so that we are close to the unlit portion of the boundary but well removed from the point of ray tangency, then the appropriate limit of (224) is a balance between the first and third terms. This leads to Airy's equation, which has already been shown to be associated with caustics. On the other hand, if we seek the solution away from the boundary but in the vicinity of the shadow boundary, then the appropriate balance is between the first two terms which, on introducing a similarity variable $\hat{n}/\sqrt{\hat{s}}$, reproduces the expected "Fresnel integral" profile.

The solution to (224)-(226) can be derived in the form

$$\hat{A}(\hat{s},\hat{n}) = -2^{1/3}\int_{-\infty}^{\infty} \frac{\mathrm{Ai}\left[-2^{1/3}v\right]\mathrm{Ai}\left[-2^{1/3}(v+\hat{n})e^{2\pi i/3}\right]e^{-i\hat{s}v}}{\mathrm{Ai}\left[-2^{1/3}ve^{2\pi i/3}\right]}\,dv, \tag{227}$$

and an asymptotic analysis of this integral is possible for all real $(\hat{s},\hat{n})$ space [38]. Here we are especially interested in the field it predicts close to the boundary but on the dark side of the point of tangency. This should fit in with Keller's theory, in which the creeping field ϕ_c is constructed as a modal expansion in the form

$$\phi_c \sim \frac{e^{ik(s+t)}}{\sqrt{t}}\sum_{m=0}^{\infty} f_m(s)e^{ik^{1/3}v_m(s)}. \tag{228}$$

The ray coordinates (s,t) are to be interpreted as follows (see Fig. 48): for an observation point O in the shadow region, construct the tangent to the scatterer passing through that point. Then t is the distance along this tangent (i.e. along the shed ray) from the body to the observation point, and s is the arc-length along the boundary between the point Q of original incident ray tangency and the point P of emission of the shed ray.

A successful match between (227) and (228) is possible provided that the functions $f_m(s)$ and $v_m(s)$ arising in (228) are given by

$$f_m(s) = \frac{\pi^{1/2}2^{1/6}e^{5\pi i/12}\mathrm{Ai}(-\lambda_m^2 e^{-2\pi i/3})}{k^{1/6}\kappa(s)^{1/3}\mathrm{Ai}'(-\lambda_m^2)}, \tag{229}$$

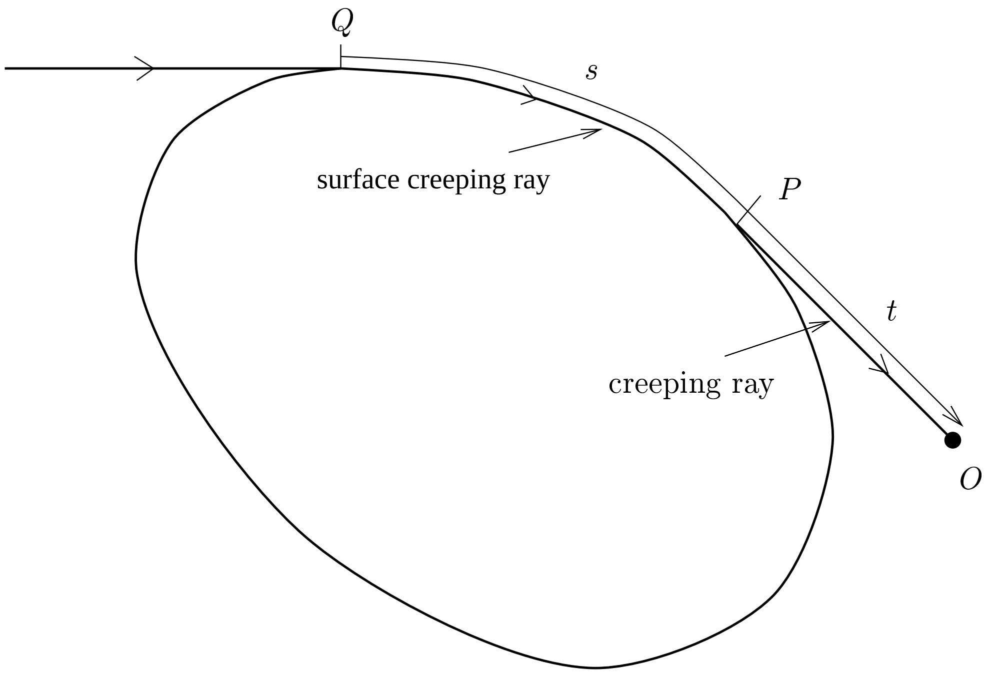


Figure 48: Diagram indicating the geometrical interpretation of creeping rays.

$$v_m(s) \quad = \quad -\lambda_m^2 e^{-2\pi i/3} 2^{-1/3} \int_0^s \kappa(q)^{2/3}\, dq, \tag{230}$$

where $\kappa(s)$ is the curvature of the body and $-\lambda_m^2$ $(m = 0, 1, 2, \cdots)$ are the zeros of the Airy function. Thus the strength of the surface creeping ray is decaying exponentially (with exponent proportional to $k^{1/3}$) as it propagates around the boundary. There is no such decay along the shed creeping rays; these rays are real rays, albeit with exponentially small magnitude.

The complex ray picture of this situation depends not so much on the amplitude (into which we can absorb the $e^{ik^{1/3}v_m(s)}$ modulation) of the creeping field, but on the phase and ray structure, since it is this that generates the Stokes surfaces responsible for switching. From (228), we see that the phase is $u = s + t$, and this can be used to generate the ray expansion of the creeping field in the form

$$x \quad = \quad x_0(s) + t x_0'(s), \tag{231}$$

$$y \quad = \quad y_0(s) + t y_0'(s), \tag{232}$$

$$u \quad = \quad s + t. \tag{233}$$

Let us first consider one of the questions posed in §3, namely the complexification of the shadow boundary. We ignore surface creeping rays that either circumnavigate the scatterer or originate at another tangency point; such rays can always be added in later. In real space the shadow boundary lies along $y = 0$, $x > 0$: for $y > 0$ the reflected and incident rays are present, while for $y < 0$ the creeping ray is present.

Let us examine the solution in the vicinity of the shadow boundary, that is, for x of order one and positive, and y small. With the ansatz $\phi = Ae^{ikx}$ we find that the correct scaling for y is $y = k^{-1/2}\bar{y}$, giving

$$\frac{\partial^2 A}{\partial \bar{y}^2} + 2i\frac{\partial A}{\partial x} = 0, \tag{234}$$

to leading order (see Fig. 47).

As we approach the shadow boundary from the outer ray region, the phases of the reflected and creeping fields are given by

$$u_r \sim x + \tfrac{y^2}{2x} + \cdots, \tag{235}$$

$$u_c \sim x + \tfrac{y^2}{2x} + \cdots. \tag{236}$$

Now we could attempt to match these fields with a similarity solution of (234) in the form $A = A(\eta)$, where $\eta = y/\sqrt{x}$, giving

$$A'' - i\eta A' = 0, \tag{237}$$

where $' \equiv d/d\eta$, and hence

$$A(\eta) = \mathrm{Fr}(\eta) = \frac{e^{-i\pi/4}}{\sqrt{2\pi}} \int_{-\infty}^{\eta} e^{iv^2/2}\, dv. \tag{238}$$

Then, as $\bar{y} \to \infty$,

$$A(x, \bar{y}) \sim 1 - \frac{ie^{-i\pi/4}}{\sqrt{2\pi}} \frac{\sqrt{x}\, e^{i\bar{y}^2/(2x)}}{\bar{y}},$$

while as $\bar{y} \to -\infty$,

$$A(x, \bar{y}) \sim -\frac{ie^{-i\pi/4}}{\sqrt{2\pi}} \frac{\sqrt{x}\, e^{i\bar{y}^2/(2x)}}{\bar{y}}.$$

Since the solution $A = 1$ represents the incident field, (238) does indeed switch off the incident field across the shadow boundary, and as $y \to \pm\infty$ the phase of A matches with the phases (235) and (236).

Unfortunately, however, the amplitude of A does not match smoothly into either the reflected or creeping fields as $|\bar{y}| \to \infty$. In fact, as $y \to 0$ the reflected field is given by

$$\phi_r \sim -\frac{\sqrt{y}}{2x\sqrt{\kappa_0}} \exp\left(\frac{iky^2}{2x} + \frac{iky^3}{8\kappa_0 x^3} + \cdots\right), \tag{239}$$

while the creeping field is given by

$$\phi_c \sim \sum_{m=0}^{\infty} \frac{f_m(0)}{\sqrt{x}} \exp\left(\frac{iky^2}{2x} - \frac{iv_m'(0)k^{1/3}y}{x} + \cdots\right). \tag{240}$$

The matching is completed by introducing a transition zone on either side of the Fresnel zone (see [38]), as shown in Fig 47. With the ansatz

$$\phi = e^{ikx} + e^{ikx} e^{iky^2/(2x)} A, \qquad y > 0, \tag{241}$$

$$\phi = e^{ikx} e^{iky^2/(2x)} A, \qquad y < 0, \tag{242}$$

we find that the correct scaling for y in the transition zones is $y = k^{-1/3}\tilde{y}$ giving

$$2\left(x\frac{\partial A}{\partial x} + \tilde{y}\frac{\partial A}{\partial \tilde{y}}\right) + A = 0, \tag{243}$$

to leading order. The solution we require is again of similarity form. We find

$$A = \frac{1}{k^{1/6}\sqrt{x}} f(\xi), \qquad \xi = \tilde{y}/x, \tag{244}$$

with

$$f \to -\frac{\sqrt{\xi}}{2\sqrt{\kappa_0}} \exp\left(\frac{i\xi^3}{8\kappa_0}\right) \qquad \text{as } \xi \to \infty, \tag{245}$$

in order to match with the reflected field (239),

$$f \to -\frac{ie^{-i\pi/4}}{\sqrt{2\pi\xi}} \qquad \text{as } \xi \to 0,$$

in order to match with the Fresnel zone, and

$$f \to \sum_{m=0}^{\infty} k^{1/6} f_m(0) e^{-iv'_m(0)\xi} \qquad \text{as } \xi \to -\infty, \tag{246}$$

in order to match with the creeping field (240). Unfortunately, the form of the solution (244) satisfies equation (243) exactly for any f, so that to determine f we need to match inwards to the Fock-Leontovich region. The details of this matching can be found in [38].

This triple layer structure of the shadow boundary introduces a new ingredient into the procedure we suggested in §4.4. Let us first consider the inner Fresnel region, in which the solution is given by (238). As $|\eta| \to \infty$ there are two rays in the approximation of (238). The first is $A \sim 1$ which matches with the incident field. The second is

$$A \sim -\frac{ie^{-i\pi/4}}{\sqrt{2\pi}} \frac{\sqrt{x}\, e^{i\bar{y}^2/(2x)}}{\bar{y}},$$

which corresponds to either the creeping or reflected field, depending on the value of $\bar{y}$, which are indistinguishable in this inner region. The Stokes line structure of A for complex η is essentially that of the error function rotated by $\pi/4$. There are active Stokes lines at $\arg\eta = 3\pi/4$ and $7\pi/4$ at which the incident field is turned off and on. The structure is shown schematically in Fig. 49.

Let us now consider the transition zones. In this case we need to determine the Stokes line structure of the function $f(\xi)$ in (244). Unfortunately, the form of f resulting from the matching with the Fock-Leontovich region is so complicated that we have found it impossible to discern the switching on and off of the incident, reflected and creeping fields by looking at its Stokes lines.

Thus we have no alternative but to resort to the more practical but less complete procedure of §4.4 which, we recall, does not require detailed knowledge of the genesis of the Stokes surface.

The position of the Stokes lines at which the incident field is switched on and off is dominated by the prefactor $e^{iky^2/(2x)}$ in (241), (242), so that these lines should pass straight through the transition regions. Here are more interested with the reflected and creeping fields, particularly since the incident field has a terminating asymptotic expansion and hence can never be responsible for switching another rayfield on or off. We need to determine the Stokes lines of the function $f(\xi)$ as $|\xi| \to \infty$. The phase of this function is $\xi^3/(8\kappa_0)$ for the reflected field, and $-v'_m(0)\xi$ for the creeping field. Hence, as $|\xi| \to \infty$ the equal-phase lines are given by $\mathrm{Re}(\xi^3) = 0$, i.e. $\arg\xi = \pi/6$, $\pi/2$, $5\pi/6$, $7\pi/6$, $3\pi/2$, $11\pi/6$, and these are the lines across which Stokes switching is possible. We know that for $\xi \to \infty$ we have only the reflected field present, while for $\xi \to -\infty$ we have only the creeping field present. Let us start from $\arg\xi = 0$ and analytically continue anti-clockwise in a large circle, as in Fig. 50. The first equal-phase line we meet is at $\arg\xi = \pi/6$. However, since the reflected field is subdominant to the creeping field on this line the creeping field cannot be switched on, and this equal-phase line cannot be a Stokes line. At the next equal phase line, $\arg\xi = \pi/2$, the reflected field is dominant, and we have the possibility that the reflected field switches on the creeping field. Indeed,

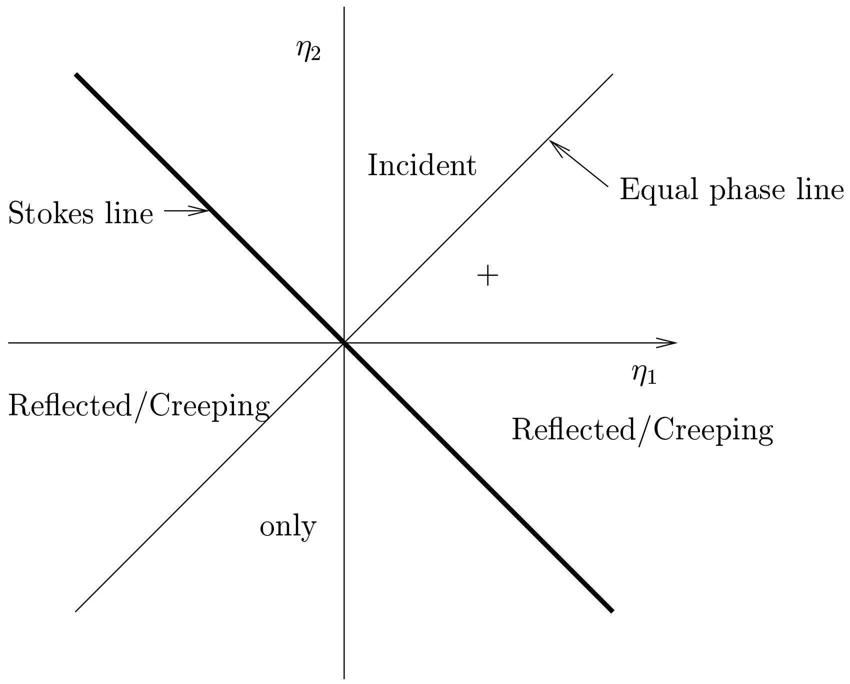


Figure 49: The Stokes structure of the reflected/creeping and incident fields in the Fresnel region of the shadow boundary. The bold lines indicate the Stokes lines, while the light lines indicate equal-phase lines which are not Stokes lines.

it must do so, since at the next equal-phase line the reflected field is subdominant, and we know that the creeping field must be present by the time we reach $\arg\xi = \pi$. Thus $\arg\xi = \pi/2$ is an active Stokes line. At the next equal-phase line the creeping field is dominant, and must turn off the reflected field, since we know that the reflected field is not present for $\arg\xi = \pi$.

Continuing into $\mathrm{Re}(\xi) < 0$, the next equal-phase line is at $\arg(\xi) = 7\pi/6$, at which the creeping field is subdominant. Thus this equal-phase line cannot be a Stokes line. Then, at the equal-phase line $\arg(\xi) = 3\pi/2$, the creeping field is dominant, and must switch on the reflected field, since this is the last chance to do so before we reach $\arg(\xi) = 2\pi$, where we know that the reflected field must be present. At the final equal-phase line, $\arg(\xi) = 11\pi/6$, the reflected field is dominant, and must switch off the creeping field since the creeping field is not present for $\arg(\xi) = 2\pi$. Thus we have been able to determine the Stokes structure for complex ξ.

When we examined the solution in the vicinity of a caustic in §4.4 we were able to check the structure derived from such an analytic continuation argument against the exact solution. Here, for reasons mentioned above, we have no exact solution to check against, but as in §4.3 we can confirm the structure shown in Fig. 50 by considering the limit

$$\lim_{n\to\infty} \arg \frac{iA_n}{A_{n-1}}$$

for the reflected field. We aim to show that as $s \to 0$ there are Stokes surfaces at $\arg\xi = \pi/2$ and $\arg\xi = 11\pi/6$. Since we will be considering the ray solution as $s \to 0$, we will now be considering the inner limit of the outer solution. Previously we have been considering the outer limit of the inner solution but, of

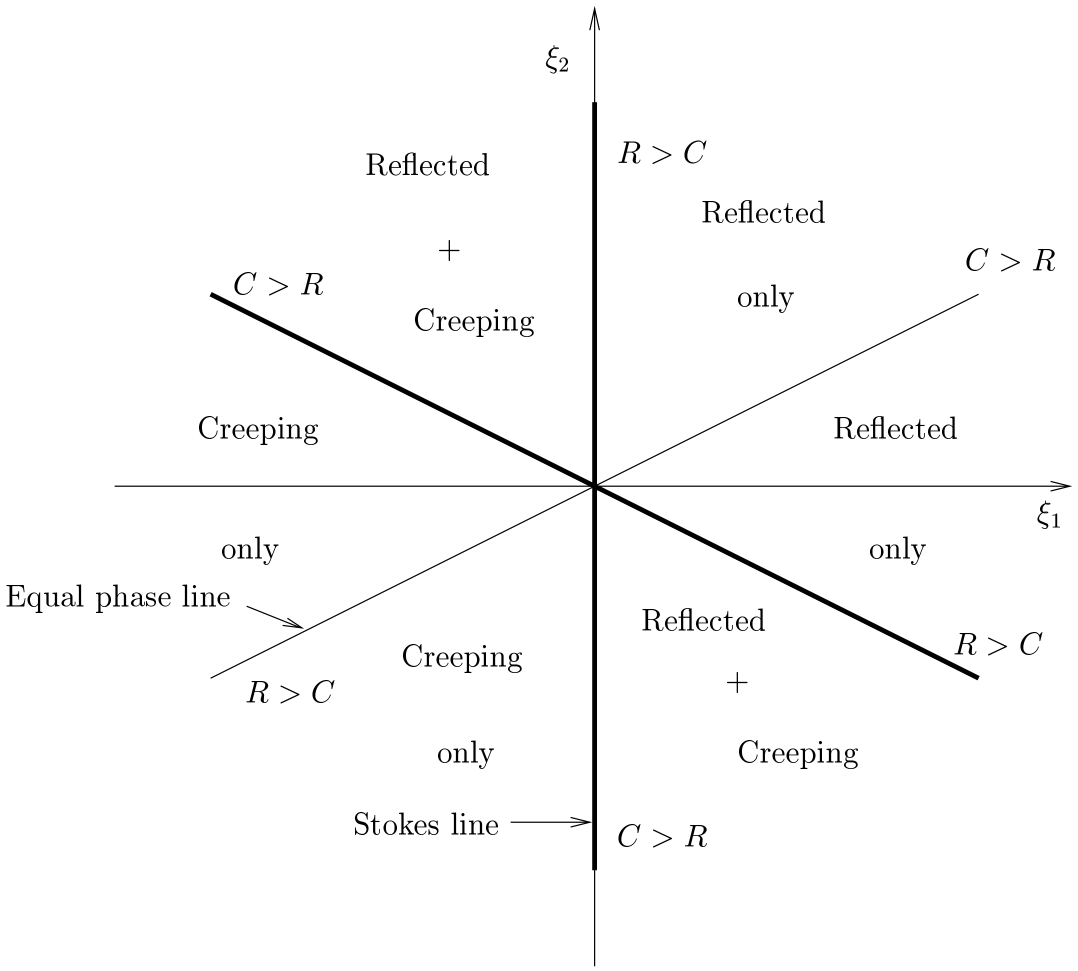


Figure 50: The Stokes structure of the reflected and creeping fields in the vicinity of the shadow boundary. The bold lines indicate the Stokes lines, while the light lines indicate equal-phase lines which are not Stokes lines. The notation $C > R$ is used to indicate lines on which the creeping field is dominant, and vice versa.

course, the two should be equivalent.

As $s \to 0$ we have, from (218)–(221),

$$\begin{aligned} x_0 &\sim s - \frac{\kappa_0^2 s^3}{6} + \cdots, && (247)\\ y_0 &\sim -\frac{\kappa_0 s^2}{2} + \cdots, && (248)\\ u_0'(s) &\sim 1 + \cdots, && (249)\\ \beta'(s) &\sim -1/2 + \cdots, && (250)\\ \alpha(s) &\sim -\frac{1}{2\kappa_0} + \cdots. && (251) \end{aligned}$$

Substituting these limits into (31) we find that

$$\begin{aligned} g_{n+1,m+1,j} &\sim -\frac{1}{32j\kappa_0^2}\left[4g''_{n,m,j-1} - 12(j-2)g'_{n,m,j-2} + 9(j-5/2)(j-1/2)g_{n,m,j-3}\right], && (252)\\ g_{n,0,0} &\sim -\sum_{m=1}^{n}\sum_{j=m}^{3m}\frac{g_{n,m,j}}{(-s/2)^j}, && (253)\\ g_{0,0,0} &\sim -(-s/2)^{1/2}. && (254) \end{aligned}$$

Thus

$$g_{n,m,j} \sim \frac{a_{n,m,j}}{(-s/2)^{3n-j-1/2}}, \tag{255}$$

where

$$\begin{aligned} a_{n+1,m+1,j} &= -\frac{1}{32j\kappa_0^2}\left[(3n-j+1/2)(3n-j+3/2)a_{n,m,j-1}\right.\\ &\qquad \left. - 6(3n-j+3/2)(j-2)a_{n,m,j-2} + 9(j-5/2)(j-1/2)a_{n,m,j-3}\right], && (256)\\ a_{n,0,0} &= -\sum_{m=1}^{n}\sum_{j=m}^{3m} a_{n,m,j}, && (257)\\ a_{0,0,0} &= -1. && (258) \end{aligned}$$

Hence

$$A_n \sim \left(-\frac{s}{2\tau}\right)^{1/2}\sum_{m=0}^{n}\sum_{j=m}^{3m}\frac{a_{n,m,j}}{(-s/2)^{3n-j}\tau^j}. \tag{259}$$

Now, as $s \to 0$ with t of order one we find

$$A_n \sim \left(-\frac{s}{2\tau}\right)^{1/2}\frac{a_{n,0,0}}{(-s/2)^{3n}}. \tag{260}$$

By calculating the first few terms in (256) we observe that $a_{n,0,0} = (-1)^n b_n$ where b_n is real and positive. Then

$$\lim_{n\to\infty} \arg\frac{iA_n}{A_{n-1}} = \arg\frac{i\Lambda}{s^3}, \tag{261}$$

where Λ is real and positive. Hence there are Stokes surfaces at which the reflected field switches on/off another rayfield wherever i/s^3 is real and positive, i.e. on $\arg(s) = \pi/6,\ 5\pi/6,\ 3\pi/2$. As $s \to 0$, $s \sim -y/(2\kappa_0 x) = -\xi/(2\kappa_0)$. Hence there will be Stokes surfaces at $\arg(\xi) = \pi/2,\ 7\pi/6,\ 11\pi/6$. However, since we are working with the tail of the WKB expansion of the reflected field, the reflected field must itself be present for these lines to be Stokes surfaces. This rules out the line $\arg(\xi) = 5\pi/6$, and we have confirmed the Stokes surfaces at which the reflected field is responsible for the Stokes switching as $\arg(\xi) = \pi/2$ and $\arg(\xi) = 11\pi/6$.

We have now given the best argument we can to identify the Stokes surfaces emanating from the shadow boundary $s \to 0, t \to 0$. Clearly this situation breaks down as we approach the tangency point $s = t = 0$. We can in principle find the Stokes surfaces near this point from the far-field analysis of (227) for complex x, y. However we have been unable to do this and we again resort to our equal phase arguments. Indeed, armed with the equal-phase surfaces for each pair of the incident, reflected and creeping rayfields, and with a knowledge of the local structure near the shadow boundary and the caustics of both the reflected and creeping fields, we hope to be able to synthesise all the components of the Stokes surface using consistency arguments, rather like building up a picture of a phase plane from a knowledge of the solution paths close to the stationary points.

The starting point is the ray solution as we emerge from the Fock-Leontovich region, so let us look locally near the tangency point by rescaling

$$s = \epsilon\kappa_0^{-2/3}\bar{s}, \qquad t = \epsilon\kappa_0^{-2/3}\bar{t} \tag{262}$$

$$x = \epsilon\kappa_0^{-2/3}\bar{x}, \qquad y = \kappa_0^{-1/3}\epsilon^2\bar{y}, \qquad u - x = \epsilon^3\bar{u}, \tag{263}$$

where $k^{-1/3} \ll \epsilon \ll 1$. Thus we are looking close to the tangency point, but not so close that we are in the inner diffraction region, and we may still use the ray solution. We first consider the reflected field, which is given by the local form of the ray equations:

$$\bar{x} = \bar{s} + \bar{t}, \tag{264}$$

$$\bar{y} = -\frac{\bar{s}^2}{2} - 2\bar{s}\bar{t}, \tag{265}$$

$$\bar{u} = 2\bar{t}\bar{s}^2, \tag{266}$$

from which we find

$$\bar{s} = \frac{2\bar{x} \mp \sqrt{4\bar{x}^2 + 6\bar{y}}}{3}, \tag{267}$$

$$\bar{t} = \frac{\bar{x} \pm \sqrt{4\bar{x}^2 + 6\bar{y}}}{3}, \tag{268}$$

$$\bar{u}_{r1} = \frac{2}{27}\left(-8\bar{x}^3 - 18\bar{x}\bar{y} + (4\bar{x}^2 + 6\bar{y})^{3/2}\right), \tag{269}$$

$$\bar{u}_{r2} = \frac{2}{27}\left(-8\bar{x}^3 - 18\bar{x}\bar{y} - (4\bar{x}^2 + 6\bar{y})^{3/2}\right). \tag{270}$$

Of the two possible solutions for the local phase, it can be shown that $\bar{u}_{r1}$ corresponds to the outgoing reflected ray, while $\bar{u}_{r2}$ corresponds to the (unphysical) incoming reflected ray. The equal-phase surfaces of $\bar{u}_{r1}$ and $\bar{u}_{r2}$ will give the Stokes surfaces for the caustic in the reflected field.

Next, in the local coordinates (262)-(263), the creeping field (228) becomes

$$\bar{x} = \bar{s} + \bar{t}, \tag{271}$$

$$\bar{y} = -\frac{\bar{s}^2}{2} - \bar{s}\bar{t}, \tag{272}$$

$$\bar{u} = \frac{\bar{s}^3}{6} + \frac{\bar{s}^2\bar{t}}{2}, \tag{273}$$

giving

$$
\begin{aligned}
\bar{s} &= \bar{x} \pm \sqrt{\bar{x}^2 + 2\bar{y}}, && (274)\\
\bar{t} &= \mp\sqrt{\bar{x}^2 + 2\bar{y}}, && (275)\\
\bar{u}_{c1} &= \frac{1}{3}\left(-\bar{x}^3 - 3\bar{x}\bar{y} + (\bar{x}^2 + 2\bar{y})^{3/2}\right), && (276)\\
\bar{u}_{c2} &= \frac{1}{3}\left(-\bar{x}^3 - 3\bar{x}\bar{y} - (\bar{x}^2 + 2\bar{y})^{3/2}\right). && (277)
\end{aligned}
$$

We see that $\bar{u}_{c1}$ corresponds to an outgoing creeping ray, propagating into the shadow zone, while $\bar{u}_{c2}$ corresponds to the other possibility of an incoming creeping ray.

The incident field is given by $\bar{u} = 0$, but we consider here only the reflected and creeping rays. As we have already mentioned, since the incident field has a terminating asymptotic expansion it cannot be responsible for the switching on or off of another rayfield, and we can therefore always add it in later once we have determined the domains in which each of the reflected and creeping fields is active.

In order to picture the Stokes structure we need to choose a two-dimensional slice of the space of complex $\bar{x}$ and $\bar{y}$. In fact we can allow three degrees of freedom if we choose $\bar{x}$ to be real and consider complex $\bar{y}$, since the equal phase surfaces depend only on $\eta = \bar{y}/\bar{x}^2$. We find that $\eta = -1/2$ corresponds to the cylinder (which is the caustic of the creeping field), $\eta = -2/3$ corresponds to the caustic of the reflected field (which is inside the cylinder), and $\eta = 0$ corresponds to the shadow boundary.

Because we now have to consider four rays and two caustics, working with a Riemann surface for the reflected field and the creeping field becomes hard to picture, and so for the present problem we label the two reflected rays u_{r1} and u_{r2} and introduce a branch cut across which u_{r1} becomes u_{r2} and vice versa. We emphasise that the branch cut is needed only because we wish to give an unambiguous name to each ray at each point, and the square root singularity in the phase means that there must be a line across which the names switch. We choose the branch cut of the reflected field to lie along the negative real axis from the caustic $\eta = -2/3$, and the branch cut of the creeping field to lie along the negative real axis from the caustic at $\eta = -1/2$.

We are now in a position to consider the equal-phase contours of each pair of rays from $\bar{u}_{r1}$, $\bar{u}_{r2}$, $\bar{u}_{c1}$, and $\bar{u}_{c2}$. These are shown in Fig. 51 for $x > 0$ and in Fig. 52 for $x < 0$.

There are three points in each of these diagrams near which we already know the local Stokes surface structure, namely the shadow boundary and the reflected- and creeping-ray caustics. For $x > 0$ the shadow boundary has the structure shown in Fig. 50, the creeping-ray caustic has the structure of an outgoing-only caustic (as shown in Fig. 19), and the reflected caustic is not present at all since neither reflected ray is present for y real and less than zero.

For $x > 0$, the rays u_{c_2} and u_{r_2} are not present near the shadow boundary $y = 0$, the creeping-ray caustic is not present since neither creeping ray is present for $x < 0$, $y/\sqrt{x} > -1/2$, and the reflected caustic has the structure of an outgoing-only caustic (again, as shown in Fig. 19).

From the knowledge of the structure in the vicinity of these points we can start to build up the global Stokes surface. A Stokes surface, once active, can only cease to be active when it meets another Stokes surface, and hence on the basis of consistency alone, the Stokes surface is as shown in Figs. 53 and 54: for $x < 0$ we can determine all the components, but for $x > 0$ we are left with more than one possibility in a number of places.

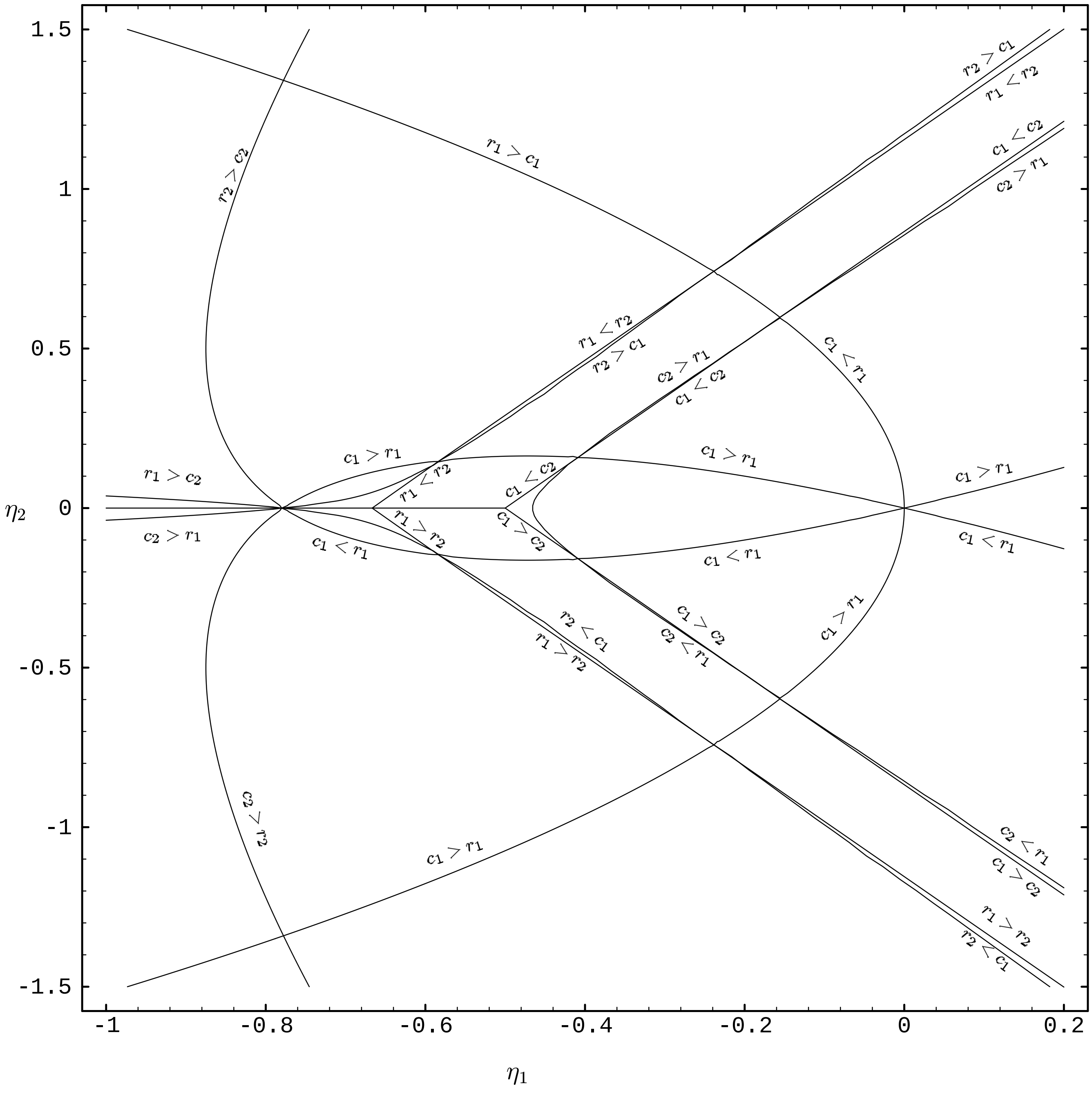


Figure 51: Equal-phase contours near the tangency point on for smooth cylindrical scatterer for $x > 0$. The figure shows the complex η plane, where $\eta = y/x^2$. The shadow boundary corresponds to $\eta = 0$, the cylinder to $\eta = -1/2$, and the caustic of the reflected field to $\eta = -2/3$.

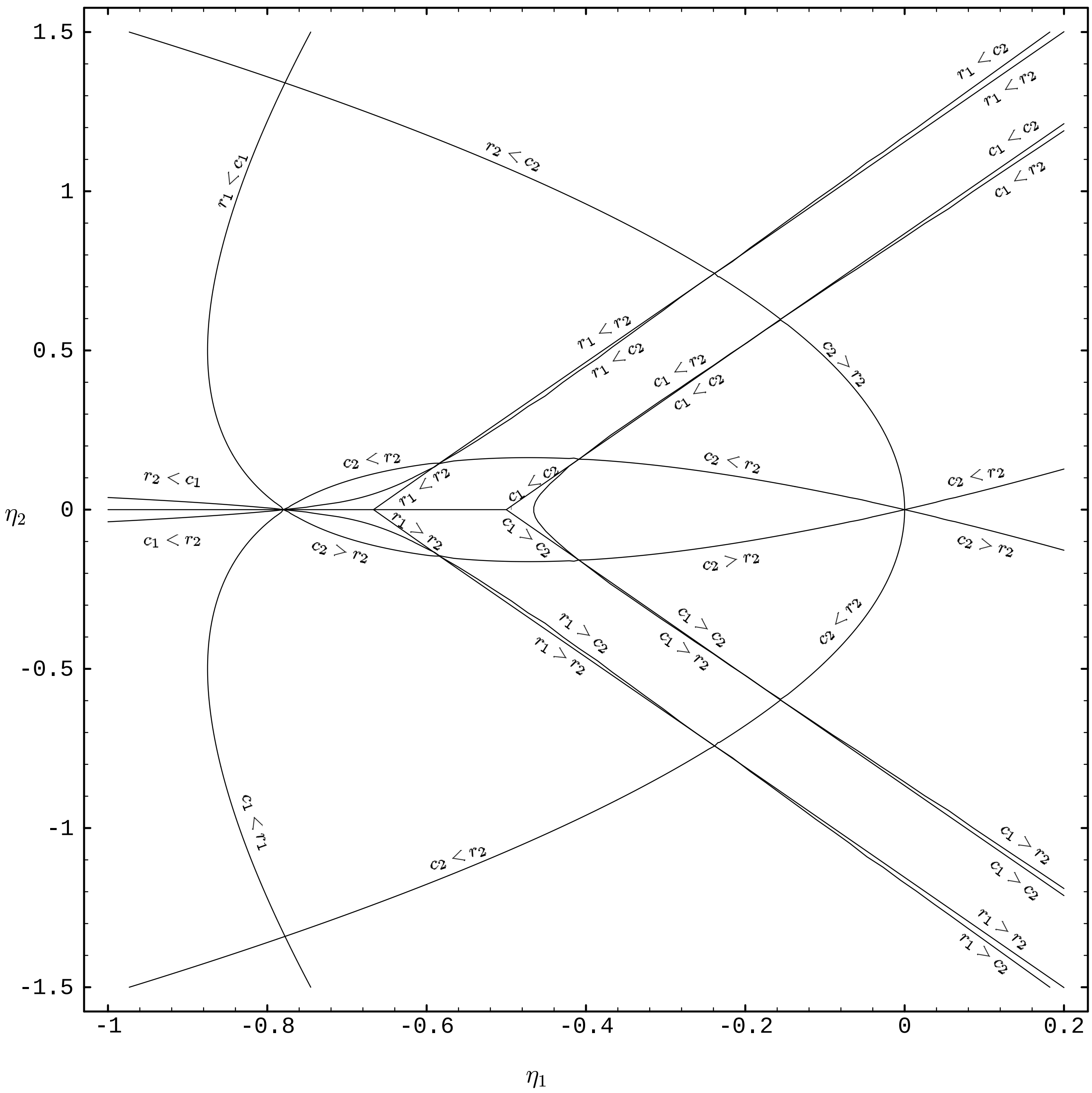


Figure 52: Equal-phase contours near the tangency point on for smooth cylindrical scatterer for $x < 0$. The figure shows the complex η plane, where $\eta = y/x^2$. The shadow boundary corresponds to $\eta = 0$, the cylinder to $\eta = -1/2$, and the caustic of the reflected field to $\eta = -2/3$.

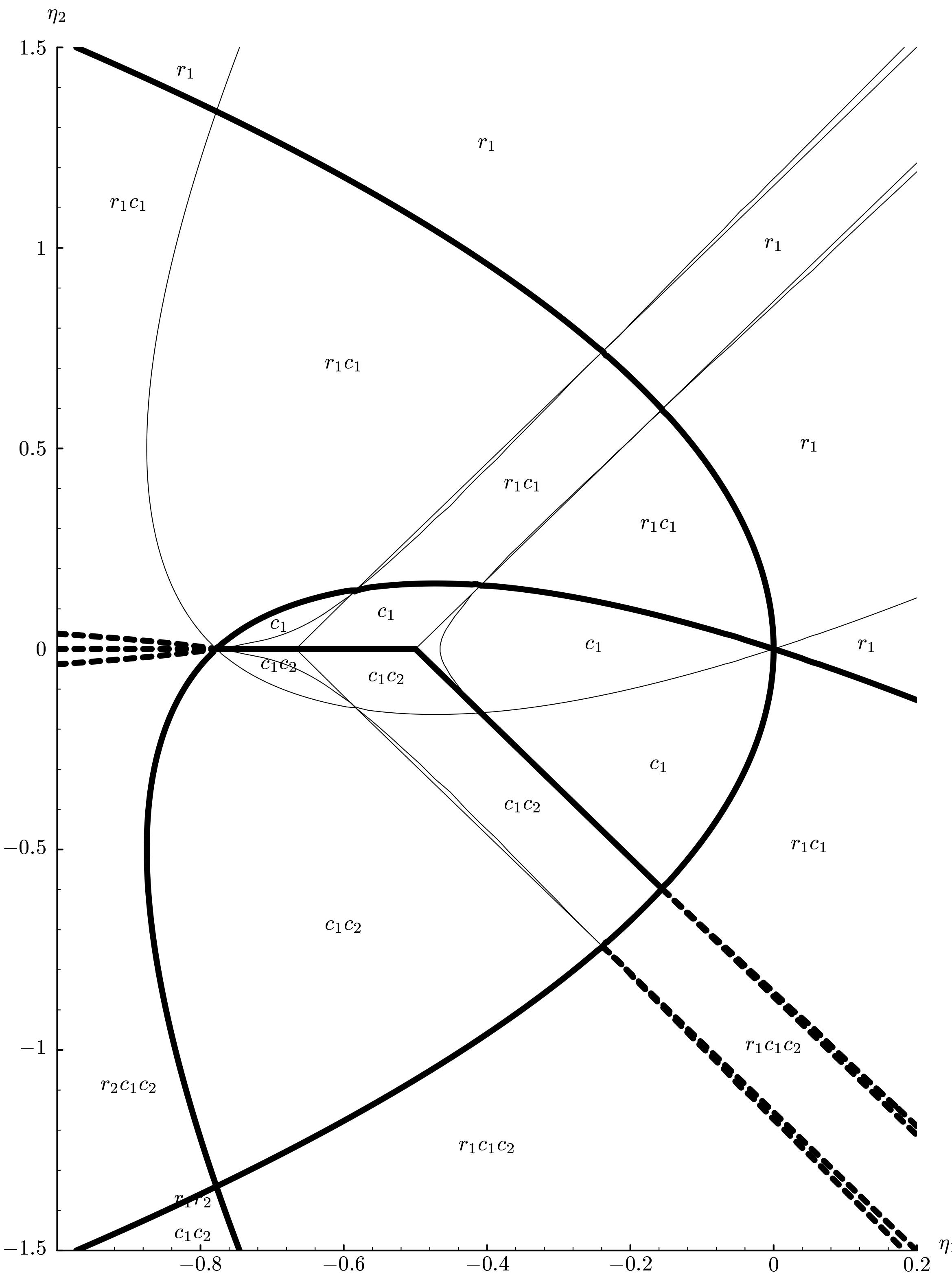


Figure 53: Stokes structure near the tangency point of a smooth cylindrical scatterer for $x > 0$. The figure shows the complex η plane, where $\eta = y/x^2$. The shadow boundary corresponds to $\eta = 0$, the cylinder to $\eta = -1/2$, and the caustic of the reflected field (which is not present for $x > 0$) to $\eta = -2/3$. The dark solid lines are active Stokes surfaces, the light solid lines are equal-phase lines which are not active Stokes surfaces, and the dashed lines are equal-phase lines for which we have insufficient information to determine whether they are active Stokes surfaces or not.

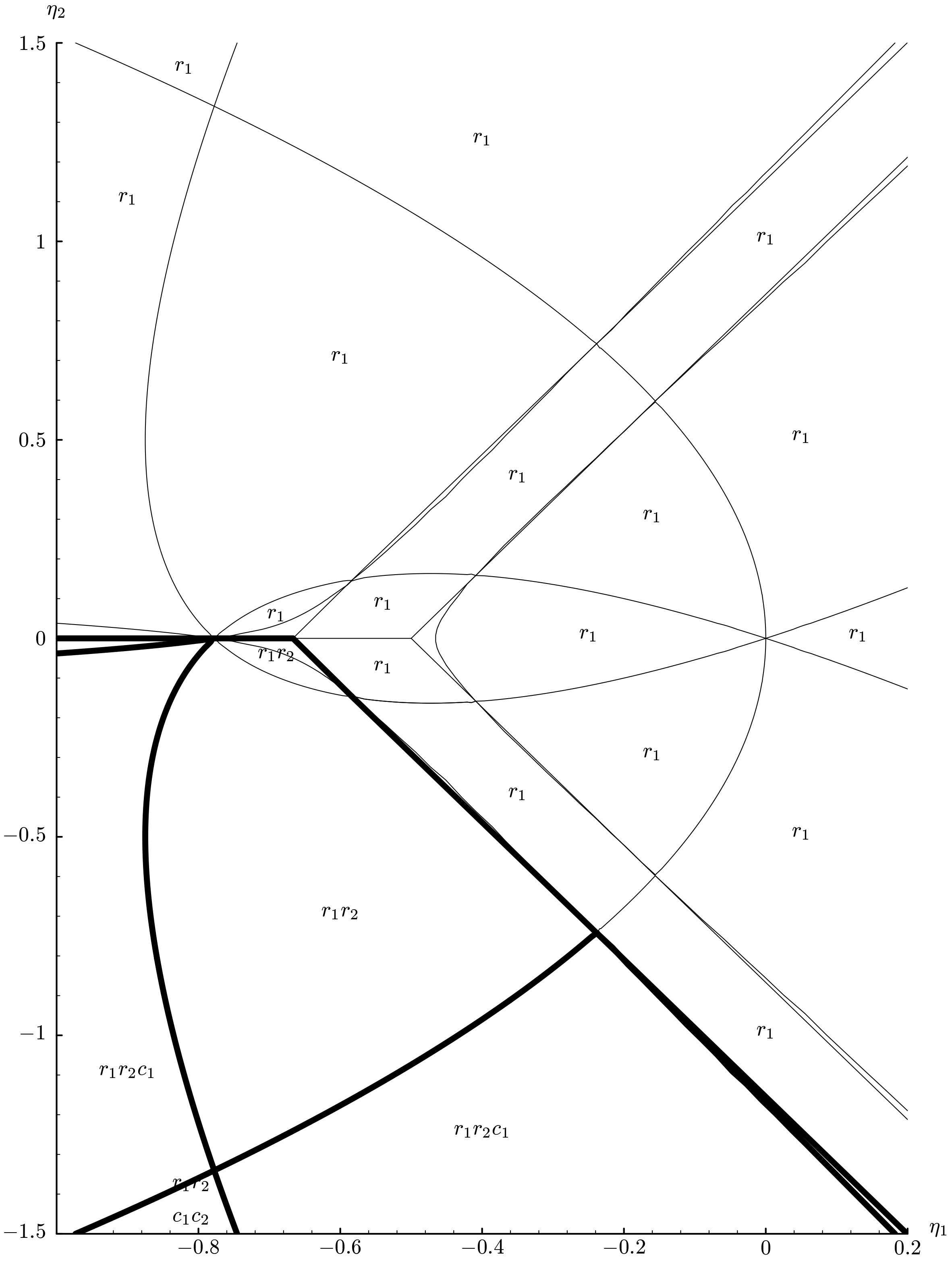


Figure 54: Stokes structure near the tangency point of a smooth cylindrical scatterer for $x < 0$. The figure shows the complex η plane, where $\eta = y/x^2$. The shadow boundary (which is not present for $x < 0$) corresponds to $\eta = 0$, the cylinder to $\eta = -1/2$, and the caustic of the reflected field to $\eta = -2/3$. The dark solid lines are active Stokes surfaces, the light solid lines are equal-phase lines which are not active Stokes surfaces.

Because no complex rays meet real space in this problem, no physical significance can be attached to Figs. 53, 54. However if the cylinder was illuminated by a *complex* plane wave, there would be complex-ray contributions to the solution in certain parts of reals space and we now examine this situation.

6.4 Scattering of a complex plane wave by a cylinder

We have seen in Figs. 53 and 54 the intricate Stokes surface which exists in complex space around the tangency point of a grazing ray in real space. Suppose, now, that the tangency point was in complex space. We might expect some of this structure to be projected into real space.

We consider the problem of a complex plane wave incident upon a cylinder. Let the incoming field be described by $\phi \sim e^{ik(x\cos a + y\sin a)}$, where a is now complex. Then the ray solution for the reflected field is given by

$$\begin{aligned} x &= x_0(s) + t\left(\left((x_0'(s))^2 - (y_0'(s))^2\right)\cos a + 2x_0'(s)y_0'(s)\sin a\right), & (278)\\ y &= y_0(s) + t\left(2x_0'(s)y_0'(s)\cos a - \left((x_0'(s))^2 - (y_0'(s))^2\right)\sin a\right), & (279)\\ u &= x_0(s)\cos a + y_0(s)\sin a + t. & (280)\end{aligned}$$

We find that $\beta(s)$ is given by

$$\beta(s) = \frac{x_0'(s)\sin a - y_0'(s)\cos a}{2(x_0'(s)y_0''(s) - y_0'(s)x_0''(s))}, \qquad (281)$$

so that there are tangency points wherever

$$x_0'(s)\sin a = y_0'(s)\cos a. \qquad (282)$$

In general, since a is complex, this will lead to complex s, and therefore complex x and y. Nevertheless, a local diffraction problem can be formulated around the tangency point, exactly as in the real case, and it leads to the formation of creeping rays in exactly the same way. To determine the domains in real space in which the incident, reflected, and creeping rays are present we need to examine the Stokes surface of the solution.

Let us assume that the imaginary part of a is small, so that the incoming wave is almost a real plane wave, and exhibits only a gradual exponential modulation in the direction $(-\sin\mathrm{Re}(a), \cos\mathrm{Re}(a))$. Then the tangency points will be close to real space, and we can perform a local analysis as before. We fix the origin of (x, y) and parametrise the cylinder as before, so that

$$x_0(0) = 0, \qquad y_0(0) = 0, \qquad x_0'(0) = 1, \qquad y_0'(0) = 0, \qquad y_0''(0) < 0.$$

If we take the real part of a to be zero then the tangency point will be close to the origin. Let us therefore set

$$\begin{aligned} x = \epsilon\kappa_0^{-2/3}\bar{x}, &\qquad y = \epsilon^2\kappa_0^{-1/3}\bar{y}, & (283)\\ a = \epsilon\kappa_0^{1/3} ib, &\qquad u - (x\cos a + y\sin a) = \epsilon^3\bar{u}. & (284)\end{aligned}$$

The ray solution (278)-(280) becomes

$$\begin{aligned} \bar{x} &= \bar{s} + \bar{t}, & (285)\\ \bar{y} &= -\frac{\bar{s}^2}{2} - 2\bar{s}\bar{t} - ib\bar{t}, & (286)\\ \bar{u} &= 2(ib + \bar{s})^2\bar{t}. & (287)\end{aligned}$$

The tangency point is at $\bar{s} = -ib$, giving $\bar{x} = -ib$, $\bar{y} = b^2/2$. Inverting (285)-(287) to determine $\bar{s}$ and $\bar{t}$ we find

$$\bar{s} = \frac{-ib + 2\bar{x}}{3} \pm \frac{(-b^2 + 2ib\bar{x} + 4\bar{x}^2 + 6\bar{y})^{1/2}}{3}, \tag{288}$$

$$\bar{t} = \frac{ib + \bar{x}}{3} \mp \frac{(-b^2 + 2ib\bar{x} + 4\bar{x}^2 + 6\bar{y})^{1/2}}{3}, \tag{289}$$

$$\bar{u}_{r1} = \frac{2}{27}\left(-ib^3 - 3b^2\bar{x} - 6ib\bar{x}^2 - 8\bar{x}^3 - 18ib\bar{y} - 18\bar{x}\bar{y} + (-b^2 + 2ib\bar{x} + 4\bar{x}^2 + 6\bar{y})^{3/2}\right), \tag{290}$$

$$\bar{u}_{r2} = \frac{2}{27}\left(-ib^3 - 3b^2\bar{x} - 6ib\bar{x}^2 - 8\bar{x}^3 - 18ib\bar{y} - 18\bar{x}\bar{y} - (-b^2 + 2ib\bar{x} + 4\bar{x}^2 + 6\bar{y})^{3/2}\right). \tag{291}$$

The ray approximation to the creeping field is given by

$$x = x_0(s) + x_0'(s)t, \tag{292}$$

$$y = y_0(s) + y_0'(s)t, \tag{293}$$

$$u = s - s_0 + x_0(s_0)\cos a + y_0(s_0)\sin a + t, \tag{294}$$

where s_0 is the value of s at tangency, that is, the solution of (282) which is close to $s = 0$ for small a. In the local coordinates (283)-(284) we find that the ray approximation becomes

$$\bar{x} = \bar{s} + \bar{t}, \tag{295}$$

$$\bar{y} = -\frac{\bar{s}^2}{2} - \bar{s}\bar{t}, \tag{296}$$

$$\bar{u} = \frac{1}{6}\left(-ib^3 + \bar{s}^3 + 3\bar{s}^2\bar{t} - 3b^2\bar{s} - 3b^2\bar{t} + 3ib\bar{s}^2 + 6ib\bar{s}\bar{t}\right). \tag{297}$$

Solving for $\bar{s}$ and $\bar{t}$ we find

$$\bar{s} = \bar{x} \pm \sqrt{\bar{x}^2 + 2\bar{y}}, \tag{298}$$

$$\bar{s} = \mp\sqrt{\bar{x}^2 + 2\bar{y}}, \tag{299}$$

$$\bar{u}_{c1} = \frac{1}{6}\left(-ib^3 - 3b^2\bar{x} - 2\bar{x}^3 - 6ib\bar{y} - 6\bar{x}\bar{y} + 2(\bar{x}^2 + 2\bar{y})^{3/2}\right), \tag{300}$$

$$\bar{u}_{c2} = \frac{1}{6}\left(-ib^3 - 3b^2\bar{x} - 2\bar{x}^3 - 6ib\bar{y} - 6\bar{x}\bar{y} - 2(\bar{x}^2 + 2\bar{y})^{3/2}\right), \tag{301}$$

We are now in a position to plot the equal-phase surfaces of the reflected, creeping and incident fields. Before we do so, let us make some further observations.

We have seen that the tangency point is $(-ib, b^2/2)$. The ray through this point, which is the "shadow boundary", is given by

$$\bar{x} = -ib + \bar{t}, \tag{302}$$

$$\bar{y} = \frac{b^2}{2} + ib\bar{t}. \tag{303}$$

This ray intersects real space in the point $x = 0$, $y = -b^2/2$, which is *inside* the cylinder (which is given by $\bar{y} = -\bar{x}^2/2$ in the local coordinates). The caustic for the reflected field is given by

$$-b^2 + 2ib\bar{x} + 4\bar{x}^2 + 6\bar{y} = 0, \tag{304}$$

and intersects real space at the point $x = 0$, $y = b^2/6$, which is *outside* the cylinder. The caustic for the creeping rays coincides with the cylinder as before. Note that although the reflected and incident rays are complex rays, the creeping rays are always real rays for a convex scatterer in two dimensions, since they must form a caustic at the boundary of the scatterer[5]. At concave parts of a boundary whispering gallery modes will be set up which will result in complex creeping rays since the exterior of the scatterer now corresponds to the interior of the creeping ray caustic.

Let us now turn to the equal-phase surfaces. As before we need to introduce branch cuts across which u_{r1} becomes u_{r2} and vice versa. We choose the branch cut for the reflected field to lie on the positive y axis from $y = b^2/6$, and define the branches so that $u_{r1} = o(x^3)$ as $x \to \infty$.

We do not need to worry about a branch cut for the creeping field, since in real space it is not possible to circumnavigate the caustic of this field. However, we do need to define the branches in the interior of the cylinder. We choose u_{c1} to be such that $(x^2 + 2y)^{3/2}$ is negative imaginary for $y < -x^2/2$.

As before, we concentrate first on the reflected and creeping fields. The equal-phase surfaces are shown in Fig. 55 for $\bar{x}$ and $\bar{y}$ real, and for $b = 1$. In Fig. 56 we show a more detailed diagram of the structure near the origin.

We now have to decide which of the equal-phase surfaces in Fig. 55 are "active" Stokes surfaces. As $b \to 0$ we expect to retrieve the solution to the problem of a real incident plane wave, with the caustic of the reflected field and the tangent ray both tending to the point $\bar{x} = 0$, $\bar{y} = 0$, and the equal-phase lines A, B, C all tending to $\bar{x} > 0$, $\bar{y} = 0$, forming the shadow boundary. As $\bar{x} \to -\infty$ we expect the reflected ray $\bar{u}_{r2}$ to be present, while as $\bar{x} \to \infty$ with $\bar{y} < 0$ we expect the creeping ray $\bar{u}_{c2}$ to be present. Thus we arrive at the picture shown in Figs. 57 and 58. The Stokes surface A intersects the cylinder at the point $\bar{x} = -\sqrt{3}\,b$ and is parallel to the tangent through this point at infinity.

Finally, we can examine the equal-phase surfaces associated with the incident field. The calculations follow precisely the procedures in (283)-(304) and we merely present the results in Figs. 59 and 60. The complete picture of active Stokes surfaces is shown in Fig. 61: it would be of interest to see if an accurate numerical simulation or a sensitive measuring device could pick up the complex wavefields, a suggestion to which we will return in the conclusion.

7 Non-analytic boundaries

One of the perceived failings of complex ray theory at present is its inability to handle non-analytic boundaries, since the first step in the solution procedure is to analytically continue the boundary and the boundary data. However, we found in §5 that the solution to Helmholtz equation in $\mathbb{C}^2$ will generically contain branch points, even when the boundary and boundary data are smooth. This gives us some hope that non-analytic boundaries can also be accounted for, and indeed we will find that when the boundary contains isolated points of non-analyticity the branch points in the solution which would be in the interior of the boundary migrate to these boundary points. Thus, in theory, piecewise analytic boundaries can be handled by complex rays.

Let us consider first the simplest problem with a non-analytic boundary, the Sommerfeld problem of the reflection of a plane wave by a half-line (see Fig. 62).

Let α be the angle the incident plane wave makes with the half-line, which is given by $y = 0$, $x < 0$,

[5]It can be shown that in three dimensions the surface creeping rays follow geodesics and may themselves form caustics on the boundary of the scatterer. In this case complex creeping rays would be shed from the dark side of these caustics.

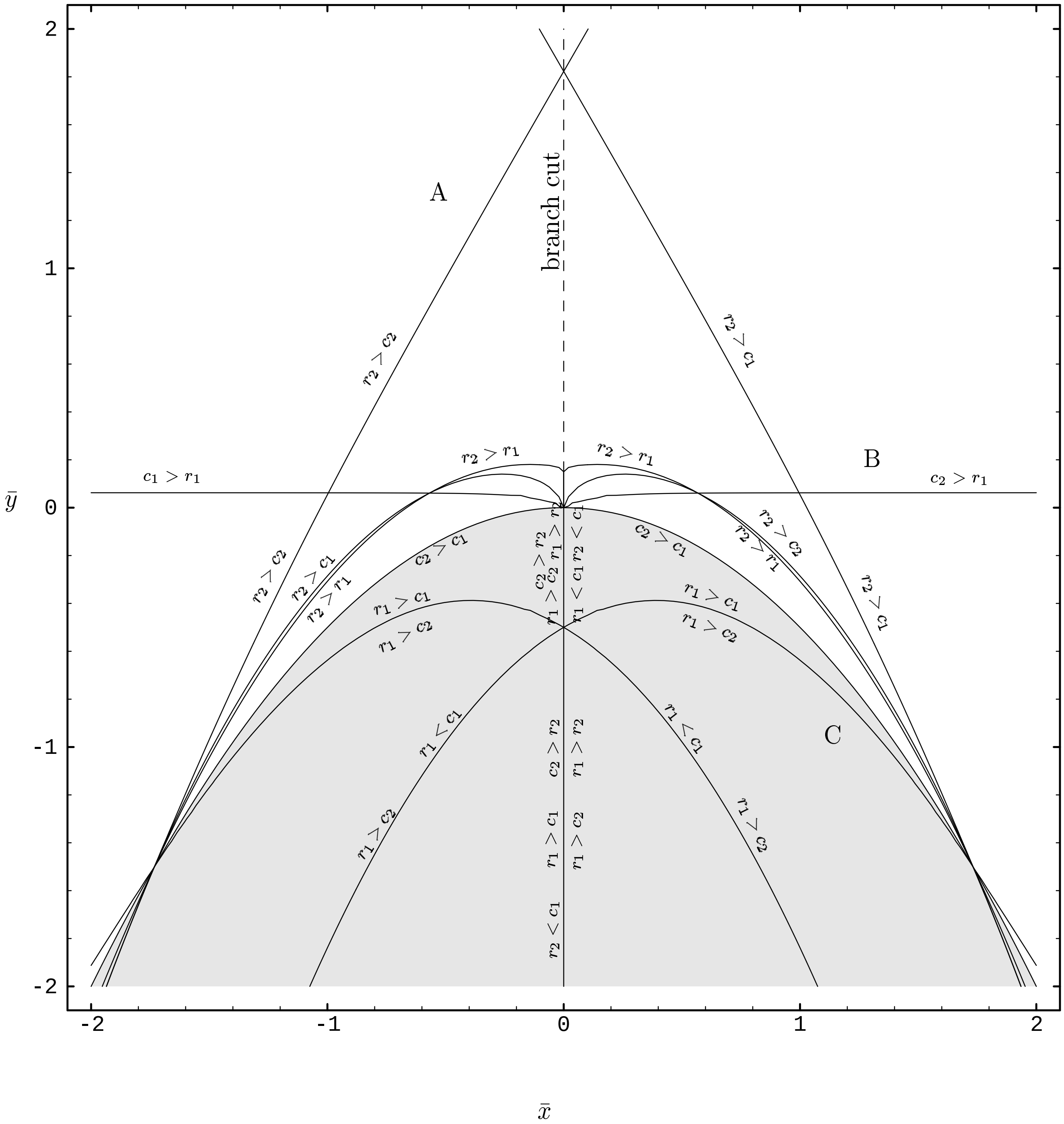


Figure 55: Equal-phase contours for the scattering of a complex plane wave (with $b = 1$) by a smooth cylinder. The figure shows real $(\bar{x}, \bar{y})$ space, and the cylinder (whose interior is shaded) corresponds to $\bar{y} = -\bar{x}^2/2$.

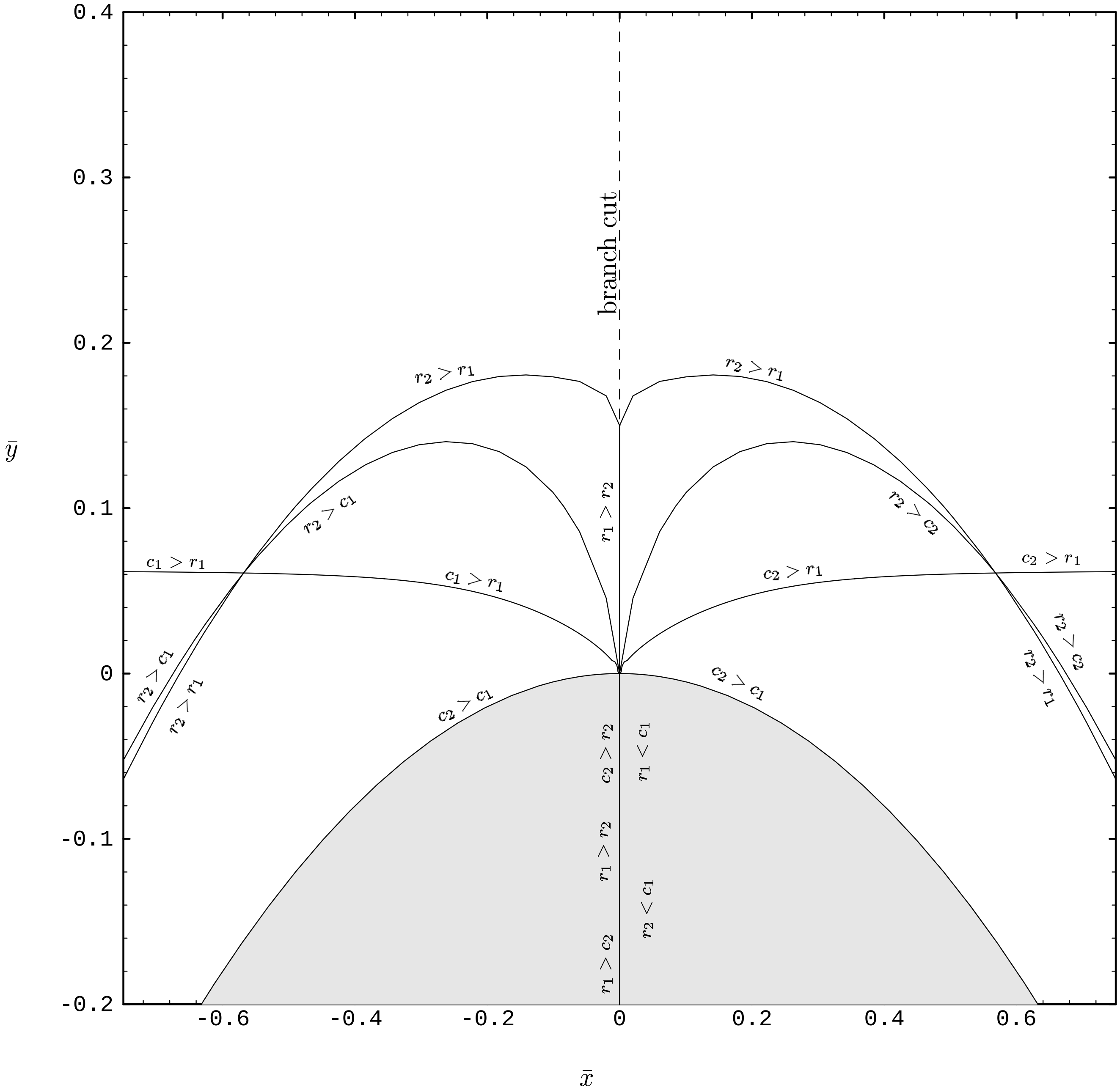


Figure 56: Enlarged diagram near $\bar{x} = \bar{y} = 0$ for the scattering of a complex plane wave (with $b = 1$) by a smooth cylinder. The figure shows real $(\bar{x}, \bar{y})$ space, and the cylinder corresponds to $\bar{y} = -\bar{x}^2/2$.

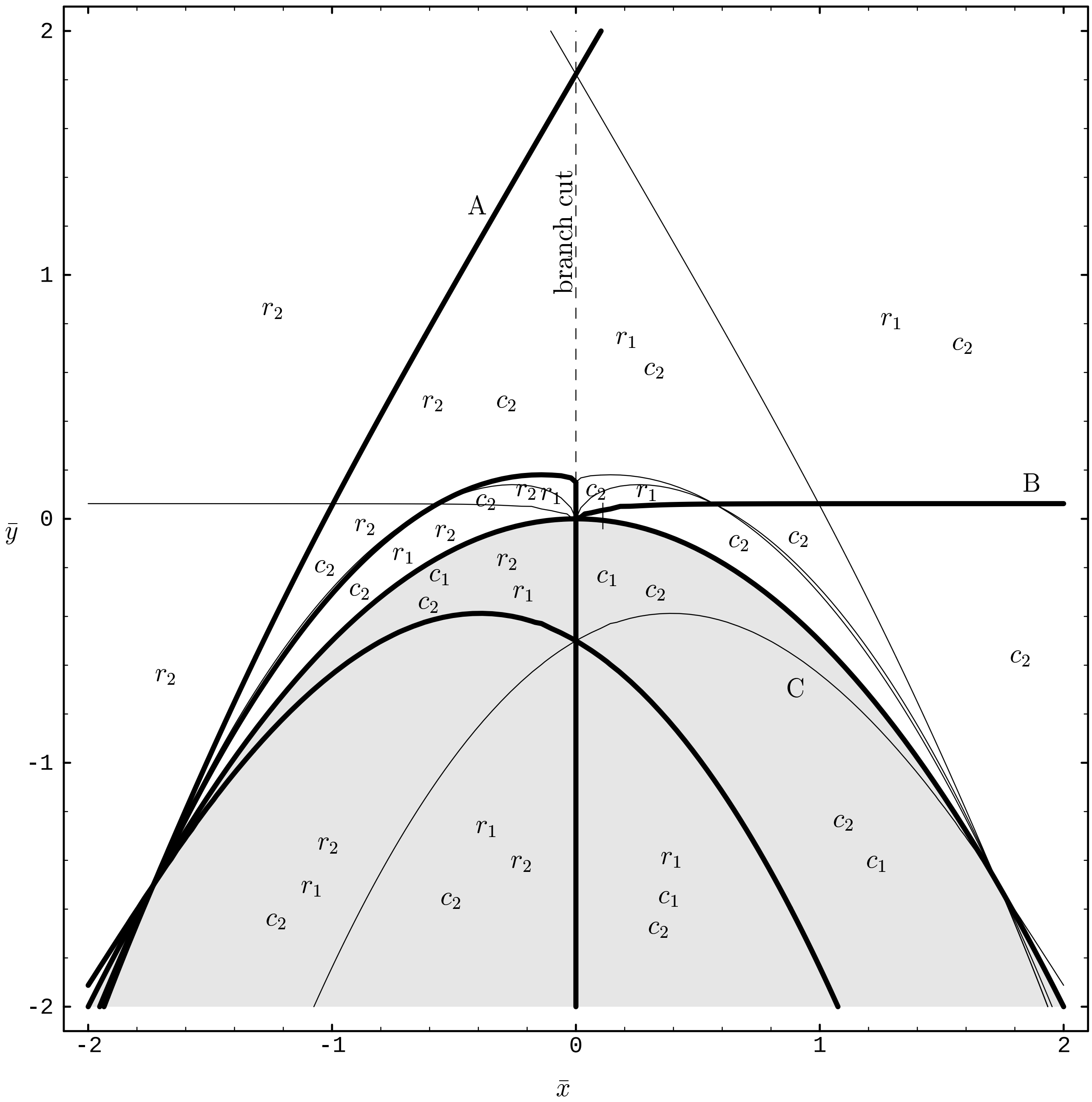


Figure 57: Reflected/creeping Stokes surfaces for the scattering of a complex plane wave (with $b = 1$) by a smooth cylinder. The figure shows real $(\bar{x}, \bar{y})$ space, and the cylinder corresponds to $\bar{y} = -\bar{x}^2/2$.

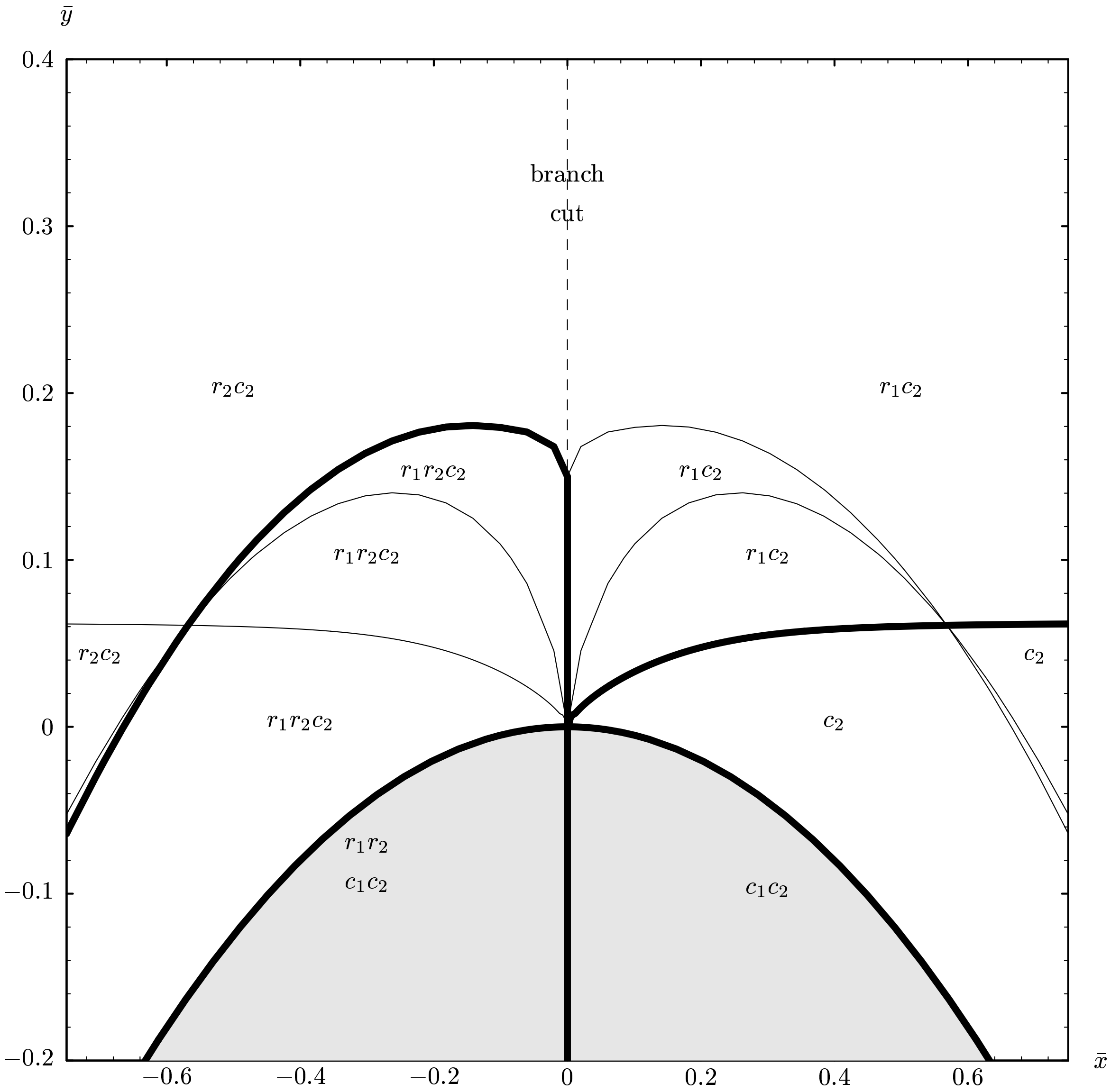


Figure 58: Enlarged diagram near $\bar{x} = \bar{y} = 0$ of the reflected/creeping Stokes surfaces for the scattering of a complex plane wave (with $b = 1$) by a smooth cylinder. The figure shows real $(\bar{x}, \bar{y})$ space, and the cylinder (interior shaded) corresponds to $\bar{y} = -\bar{x}^2/2$.

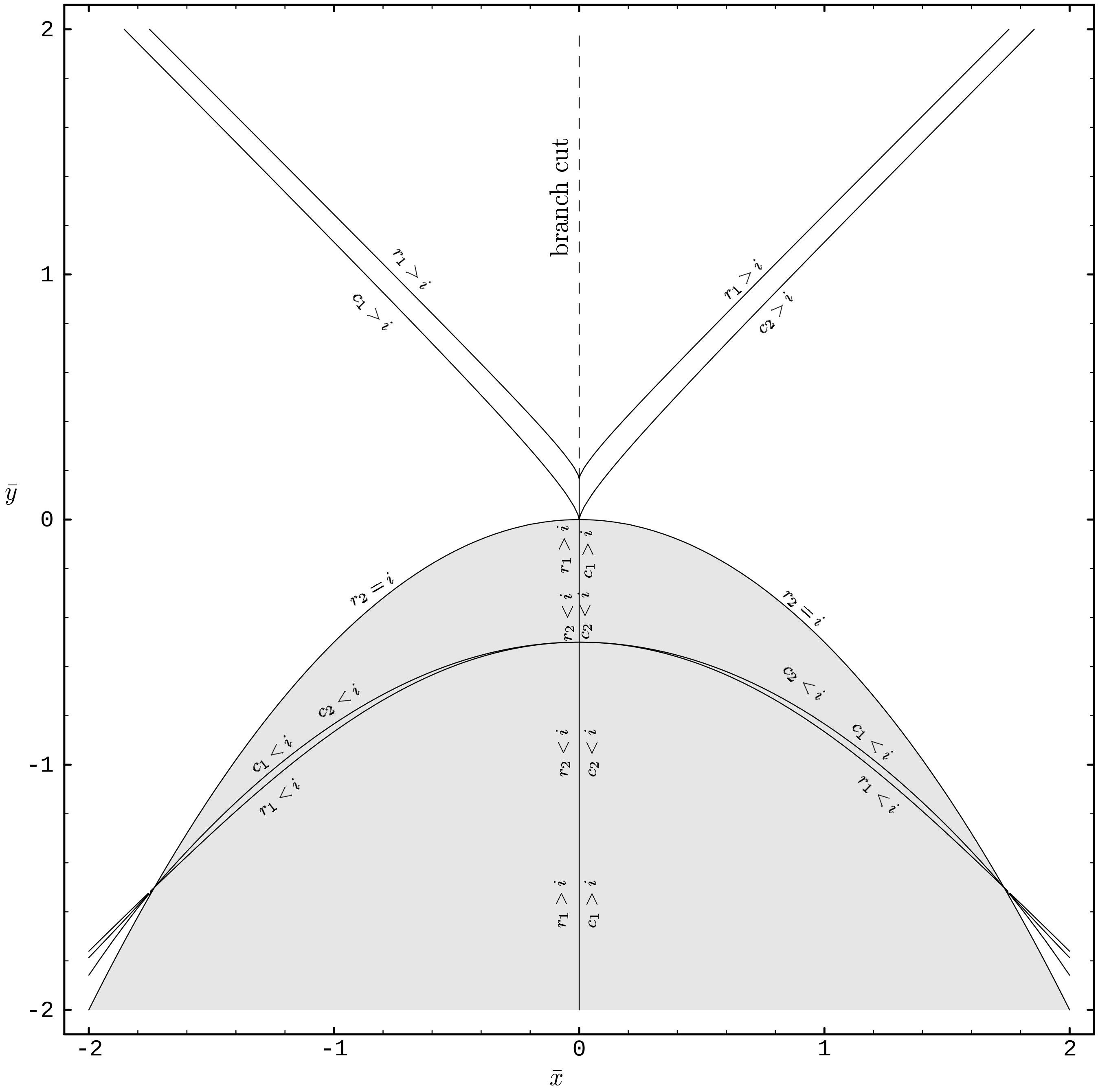


Figure 59: Equal phase lines for incident and either reflected or creeping fields for the scattering of a complex plane wave (with $b = 1$) by a smooth cylinder. The figure shows real $(\bar{x}, \bar{y})$ space, and the cylinder corresponds to $\bar{y} = -\bar{x}^2/2$.

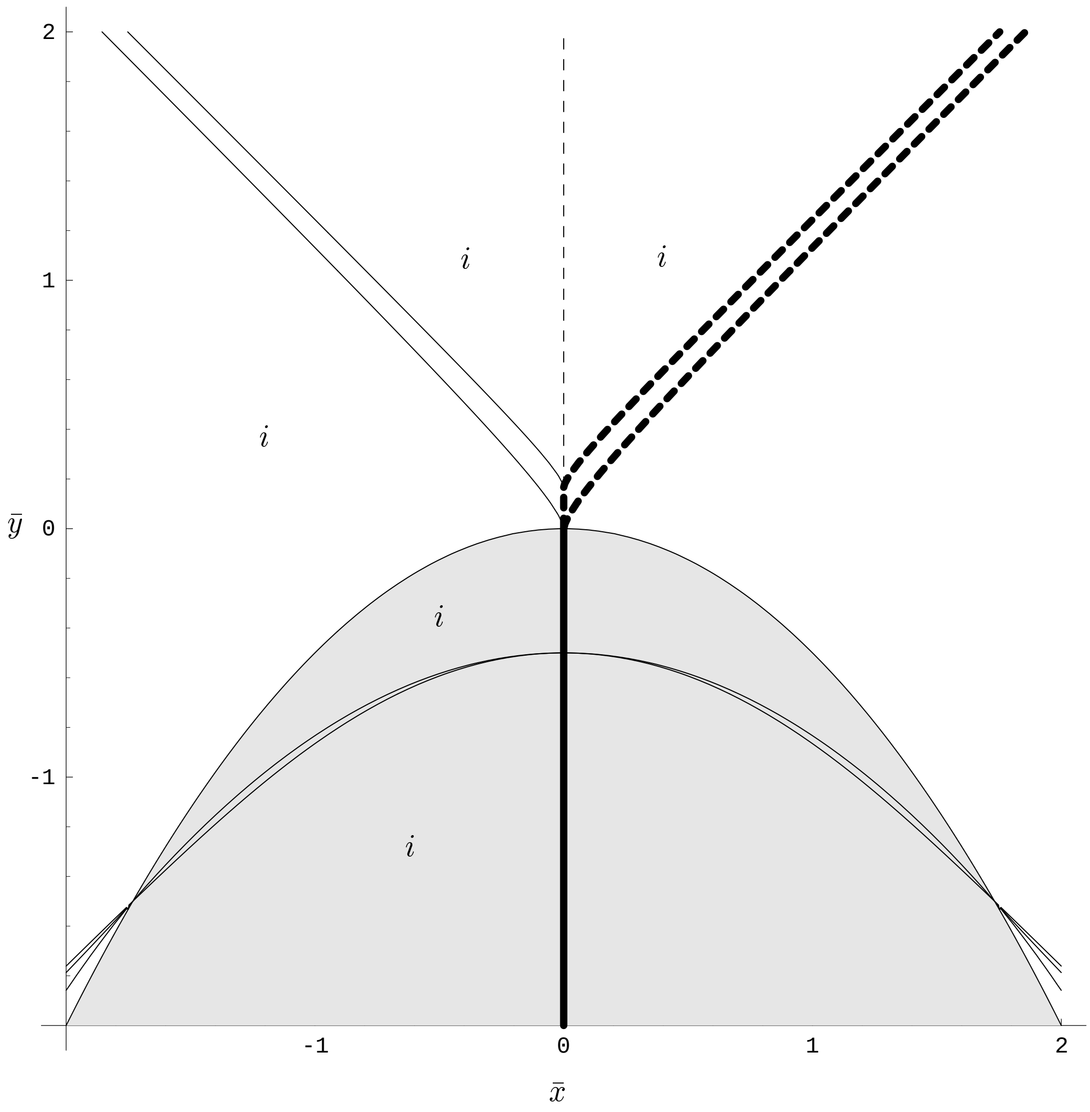


Figure 60: Stokes surfaces for the incident field for the scattering of a complex plane wave (with $b = 1$) by a smooth cylinder. The figure shows real $(\bar{x}, \bar{y})$ space, and the cylinder (interior shaded) corresponds to $\bar{y} = -\bar{x}^2/2$. We are unable to determine which one of the dark dashed lines is a Stokes line and which is an inactive equal phase line; this depends on whether the incident field is turned off by the reflected field r_1 or the creeping field c_2.

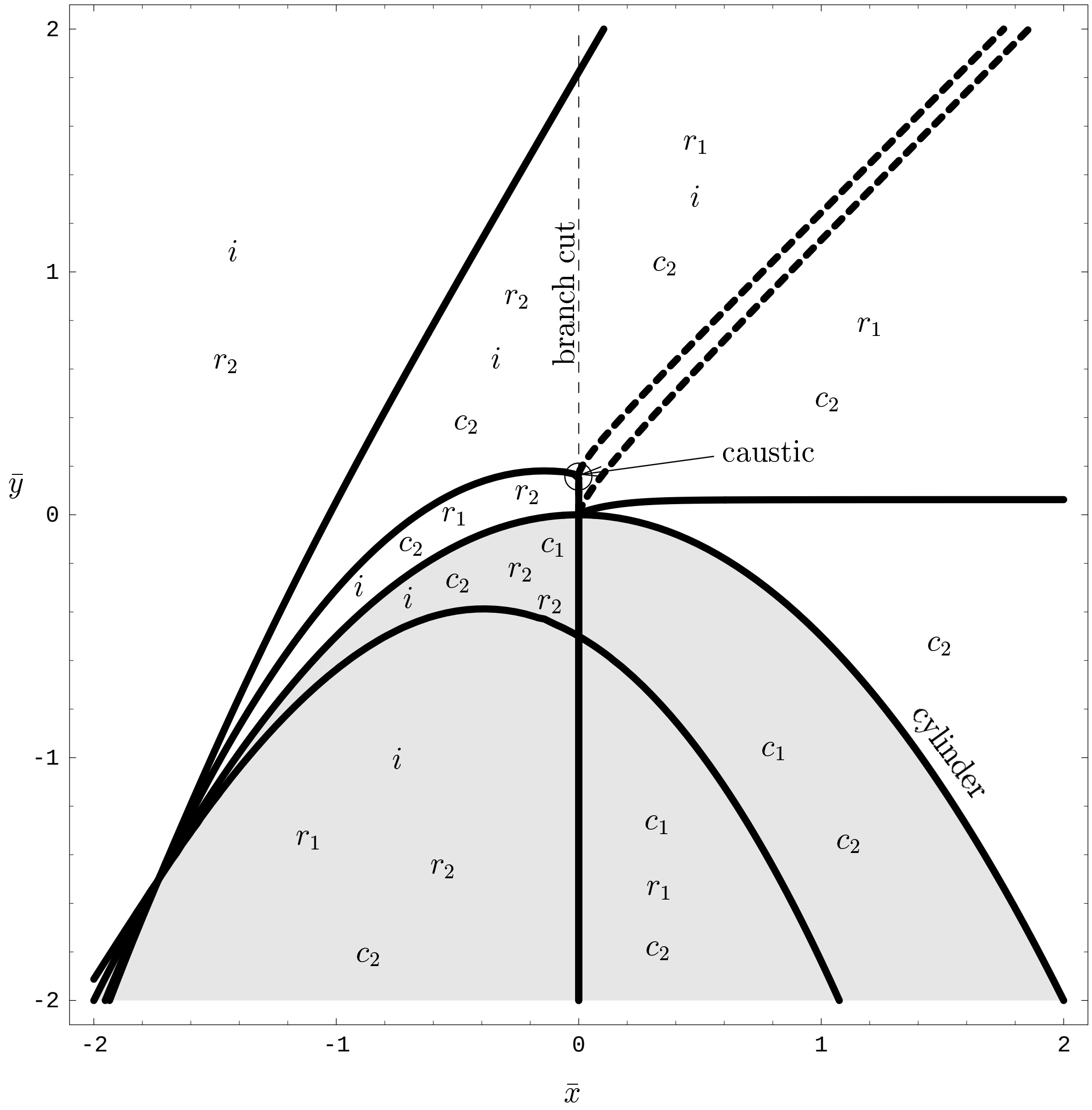


Figure 61: Complete Stokes structure for the scattering of a complex plane wave (with $b = 1$) by a smooth cylinder. The figure shows real $(\bar{x}, \bar{y})$ space, and the cylinder (interior shaded) corresponds to $\bar{y} = -\bar{x}^2/2$. The only uncertainty is which of the dashed lines is responsible for switching off the incident field.

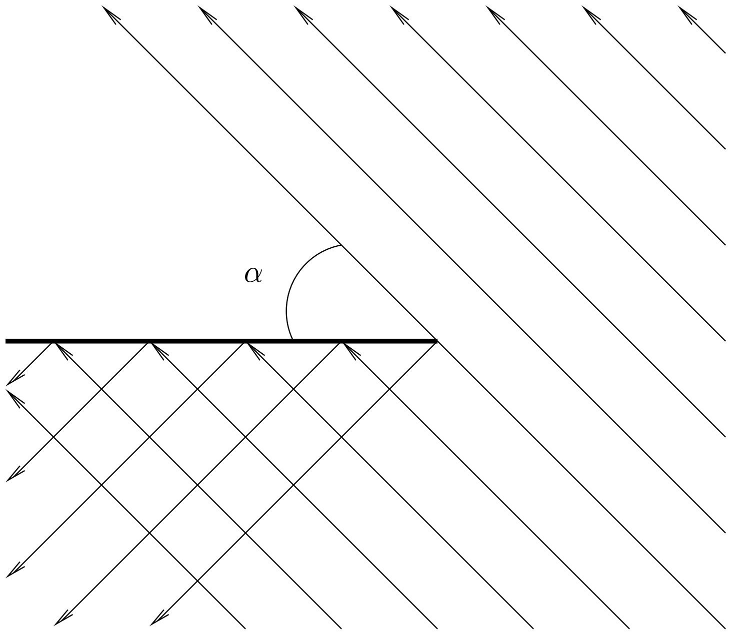


Figure 62: The Sommerfeld Problem of plane wave incidence on a reflecting half-line.

as shown in Fig. 62. Then the exact solution may be obtained by Fourier transform and Wiener–Hopf factorisation, or by the Kantorovich–Lebedev transform [17, 18], and is given by

$$\phi = \frac{e^{-i\pi/4+ikr}}{\sqrt{\pi}}\left[-\,\mathrm{Fr}\left(\sqrt{2kr}\cos\left(\frac{\theta-\alpha}{2}\right)\right)+\mathrm{Fr}\left(-\sqrt{2kr}\cos\left(\frac{\theta+\alpha}{2}\right)\right)\right], \tag{305}$$

where

$$\mathrm{Fr}(z) = e^{-iz^2}\int_z^\infty e^{iv^2}\,dv, \tag{306}$$

is the Fresnel integral. The ray approximation corresponds to the far field $r \to \infty$ (since there is no length scale in the problem), and is given by

$$\begin{aligned}\phi &\sim e^{-ikr\cos(\theta+\alpha)} - e^{-ikr\cos(\theta-\alpha)}\\ &\quad -\frac{e^{ikr+i\pi/4}}{2\sqrt{2\pi kr}}\left(\frac{1}{\cos((\theta+\alpha)/2)}+\frac{1}{\cos((\theta-\alpha)/2)}\right), \qquad -\pi<\theta<-\pi+\alpha,\end{aligned} \tag{307}$$

$$\begin{aligned}\phi &\sim e^{-ikr\cos(\theta+\alpha)}\\ &\quad -\frac{e^{ikr+i\pi/4}}{2\sqrt{2\pi kr}}\left(\frac{1}{\cos((\theta+\alpha)/2)}+\frac{1}{\cos((\theta-\alpha)/2)}\right), \qquad -\pi+\alpha<\theta<\pi-\alpha,\end{aligned} \tag{308}$$

$$\phi \sim -\frac{e^{ikr+i\pi/4}}{2\sqrt{2\pi kr}}\left(\frac{1}{\cos((\theta+\alpha)/2)}+\frac{1}{\cos((\theta-\alpha)/2)}\right), \qquad \pi-\alpha<\theta<\pi. \tag{309}$$

We see that there are three rayfields; the incident (I),

$$u = -r\cos(\theta+\alpha), \tag{310}$$

$$A = 1, \tag{311}$$

the reflected (R),

$$u = -r\cos(\theta-\alpha), \tag{312}$$
$$A = -1, \tag{313}$$

and a diffracted wavefield originating at the tip of the half-line (D),

$$u = r, \tag{314}$$
$$A = -\frac{e^{i\pi/4}}{2\sqrt{2\pi kr}}\left(\frac{1}{\cos((\theta+\alpha)/2)}+\frac{1}{\cos((\theta-\alpha)/2)}\right). \tag{315}$$

In our study of the circular caustic problem in §4.4, we found it useful to analytically continue the exact solution in the inner region and interpret it in terms of complex rays. Here, since we have the new problem of non-analyticity, we adopt a similar approach by analytically continuing the exact solution (305). We see that the tip of the half-line $r = 0$ is a branch point of ϕ, and that the branch chosen corresponds to $-\pi < \theta < \pi$. Analytically continuing with r real we find a second branch defined by $\pi < \theta < 3\pi$ with the ray approximation

$$\phi \sim -\frac{e^{ikr+i\pi/4}}{2\sqrt{2\pi kr}}\left(\frac{1}{\cos((\theta+\alpha)/2)}+\frac{1}{\cos((\theta-\alpha)/2)}\right), \qquad \pi<\theta<\pi+\alpha, \tag{316}$$

$$\phi \sim -e^{-ikr\cos(\theta-\alpha)} - \frac{e^{ikr+i\pi/4}}{2\sqrt{2\pi kr}}\left(\frac{1}{\cos((\theta+\alpha)/2)}+\frac{1}{\cos((\theta-\alpha)/2)}\right), \qquad \pi+\alpha<\theta<3\pi-\alpha, \tag{317}$$

$$\phi \sim e^{-ikr\cos(\theta+\alpha)} - e^{-ikr\cos(\theta-\alpha)} - \frac{e^{ikr+i\pi/4}}{2\sqrt{2\pi kr}}\left(\frac{1}{\cos((\theta+\alpha)/2)}+\frac{1}{\cos((\theta-\alpha)/2)}\right), \qquad 3\pi-\alpha<\theta<3\pi, \tag{318}$$

which is illustrated in Fig. 63.

The reflected wave of the first problem is the incident wave of the second, and vice versa. Clearly the solution (305) has period 4π in θ, and therefore for real x and y lives on a two-sheeted Riemann surface[6]. The branch point is in real space, and the branch cut must lie outside physical space, and is therefore constrained to lie along the reflecting half-line.

We note that when we complexify x and y, there are in fact 8 sheets to the Riemann surface.

Let us examine the Stokes surface structure of the solution (305). The Stokes surfaces are simply the Stokes surfaces of the Fresnel integral, $\mathrm{Fr}(z)$, and are therefore at $\arg z = 3\pi/4$, $7\pi/4$; there are equal-phase surfaces which are not Stokes surfaces at $\arg z = \pi/4$, $5\pi/4$. Hence the Stokes surfaces are given by

$$\cos\left(\frac{\theta_1\pm\alpha}{2}\right)\cosh\left(\frac{\theta_2}{2}\right)-\sin\left(\frac{\theta_1\pm\alpha}{2}\right)\sinh\left(\frac{\theta_2}{2}\right)=0 \tag{319}$$

while the remaining equal-phase surfaces are given by

$$\cos\left(\frac{\theta_1\pm\alpha}{2}\right)\cosh\left(\frac{\theta_2}{2}\right)+\sin\left(\frac{\theta_1\pm\alpha}{2}\right)\sinh\left(\frac{\theta_2}{2}\right)=0, \tag{320}$$

where $\theta = \theta_1 + i\theta_2$, θ_1, θ_2 real. These are shown in Fig. 64. The Stokes surfaces intersect real space in the shadow boundaries $\theta = -\pi+\alpha$ and $\theta = \pi-\alpha$.

[6]Indeed, Sommerfeld himself solved the problem by recognising that the solution should have period 4π in θ, removing the half-line, and imposing suitable radiation conditions on the Riemann surface $-\pi < \theta < 3\pi$ [40, 3].

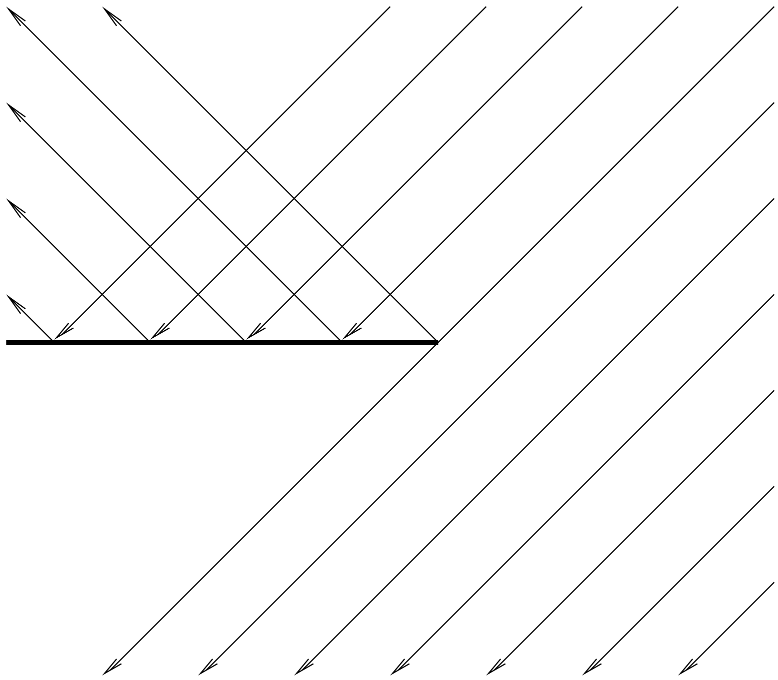


Figure 63: The Dual Sommerfeld Problem.

We see that at each Stokes surface the diffracted field D turns on or off one of the incident or reflected fields, I and R. A pair of equal-phase surfaces emerges from each shadow boundary, one of which is an active Stokes surface and one of which is inactive.

The same kind of analysis can be carried out for wedges of arbitrary angle and, indeed, these "Sommerfeld" problems are the local inner diffraction problems for arbitrary boundaries having "ends" when the ray which hits the singularity in the boundary is real. Since we are hoping to generalise the Geometrical Theory of Diffraction to complex rays, let us now consider the inner diffraction problem when the ray which hits the singular point on the boundary is a complex ray. This problem has been considered in [6], and corresponds simply to allowing the angle of incidence α to be complex, with the solution (305) still valid.

Since α is complex, the "origin" of the Stokes surfaces is in complex space. The active Stokes surfaces are given by

$$\cos\left(\frac{\theta_1 \pm \alpha_1}{2}\right)\cosh\left(\frac{\theta_2 \pm \alpha_2}{2}\right) - \sin\left(\frac{\theta_1 \pm \alpha_1}{2}\right)\sinh\left(\frac{\theta_2 \pm \alpha_2}{2}\right) = 0 \tag{321}$$

while the inactive equal-phase surfaces are given by

$$\cos\left(\frac{\theta_1 \pm \alpha_1}{2}\right)\cosh\left(\frac{\theta_2 \pm \alpha_2}{2}\right) + \sin\left(\frac{\theta_1 \pm \alpha_1}{2}\right)\sinh\left(\frac{\theta_2 \pm \alpha_2}{2}\right) = 0, \tag{322}$$

where $\alpha = \alpha_1 + i\alpha_2$, α_1, α_2 real.

The intersection with real space gives the ray approximation

$$\begin{aligned} \phi \quad \sim \quad & -e^{-ikr\cos(\theta+\alpha)} + e^{-ikr\cos(\theta-\alpha)} \\ & + \frac{e^{ikr-i\pi/4}}{2\sqrt{2\pi kr}}\left(\frac{1}{\cos((\theta+\alpha)/2)} - \frac{1}{\cos((\theta-\alpha)/2)}\right), -\pi < \theta < \theta^*, \end{aligned} \tag{323}$$

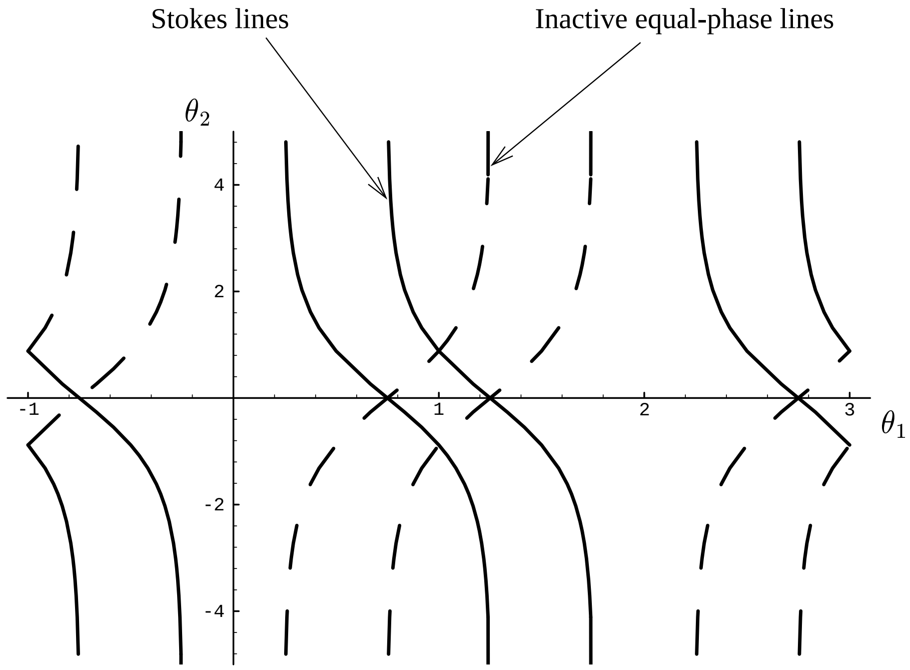


Figure 64: Stokes surfaces of the Sommerfeld problem with $\alpha = \pi/4$ where $\theta = \theta_1 + i\theta_2$, θ_1, $\theta_2 \in \mathbb{R}$.

$$\phi \sim -e^{-ikr\cos(\theta+\alpha)} + \frac{e^{ikr-i\pi/4}}{2\sqrt{2\pi kr}}\left(\frac{1}{\cos((\theta+\alpha)/2)} - \frac{1}{\cos((\theta-\alpha)/2)}\right), \theta^* < \theta < \theta^{**}, \tag{324}$$

$$\phi \sim \frac{e^{ikr-i\pi/4}}{2\sqrt{2\pi kr}}\left(\frac{1}{\cos((\theta+\alpha)/2)} - \frac{1}{\cos((\theta-\alpha)/2)}\right), \theta^{**} < \theta < \pi, \tag{325}$$

where θ^* and θ^{**} are the solutions of

$$\cos\left(\frac{\theta^*+\alpha_1}{2}\right)\cosh\left(\frac{\alpha_2}{2}\right) - \sin\left(\frac{\theta^*+\alpha_1}{2}\right)\sinh\left(\frac{\alpha_2}{2}\right) = 0, \tag{326}$$

$$\cos\left(\frac{\theta^{**}-\alpha_1}{2}\right)\cosh\left(\frac{\alpha_2}{2}\right) + \sin\left(\frac{\theta^{**}-\alpha_1}{2}\right)\sinh\left(\frac{\alpha_2}{2}\right) = 0, \tag{327}$$

namely

$$\theta^* = 2\tan^{-1}\left[\frac{\cos(\alpha_1/2)\cosh(\alpha_2/2) - \sin(\alpha_1/2)\sinh(\alpha_2/2)}{\sin(\alpha_1/2)\cosh(\alpha_2/2) + \cos(\alpha_1/2)\sinh(\alpha_2/2)}\right], \tag{328}$$

$$\theta^{**} = 2\tan^{-1}\left[\frac{-\cos(\alpha_1/2)\cosh(\alpha_2/2) + \sin(\alpha_1/2)\sinh(\alpha_2/2)}{\sin(\alpha_1/2)\cosh(\alpha_2/2) + \cos(\alpha_1/2)\sinh(\alpha_2/2)}\right]. \tag{329}$$

As before there are two shadow boundaries, but the relationship between the direction of the shadow boundaries and the complex angle of incidence is now given by (328) and (329). By equating the real parts of the diffracted and the incident/reflected fields we find that (328), (329) may be written as

$$\theta^* = -\alpha_1 + \cos^{-1}\left(\frac{-1}{\cosh\alpha_2}\right), \tag{330}$$

$$\theta^{**} = \alpha_1 + \cos^{-1}\left(\frac{-1}{\cosh\alpha_2}\right), \tag{331}$$

where the branch of $\cos^{-1}$ is chosen so that the imaginary part of the incident/reflected phase is larger than the imaginary part of the diffracted phase (that is, the diffracted field is dominant). Thus we see that the shadow boundaries may be rotated by up to $\pi/2$ in either direction depending on the value of α_2. This is clear also from Fig. 64. Changing the value of α_2 corresponds to moving the origin of the Stokes surfaces up or down in this figure, as seen is Figs. 65-68. Clearly the intersection of the Stokes surfaces with real space will then change by up to an angle of $\pi/2$. In the right-hand sides of Figs. 65-68 we show the behaviour of the ray solution in real space, where the arrows indicate the normals to the level curves of $\mathrm{Re}(u)$, which would be the ray directions, were the rays real. These are also the level curves of $\mathrm{Im}(u)$, and thus indicate lines of constant magnitude of the wavefield.

Note that in Fig. 66b the Stokes surfaces have actually crossed the branch cut, so that the ones that were active in Fig. 66a have moved up onto the second branch, while those on the second branch have moved down onto the first branch.

Again, it would be interesting to know if these predictions could be detected by a sensitive measuring device.

7.1 Gaussian beam incident on a half-line

As an example of the complex GTD let us now consider the problem of a Gaussian Beam incident on a half-line.

The methodology of the traditional GTD is to calculate the incident and reflected fields, perform an inner asymptotic analysis in the vicinity of the singular point, calculate the far field of this inner problem, and

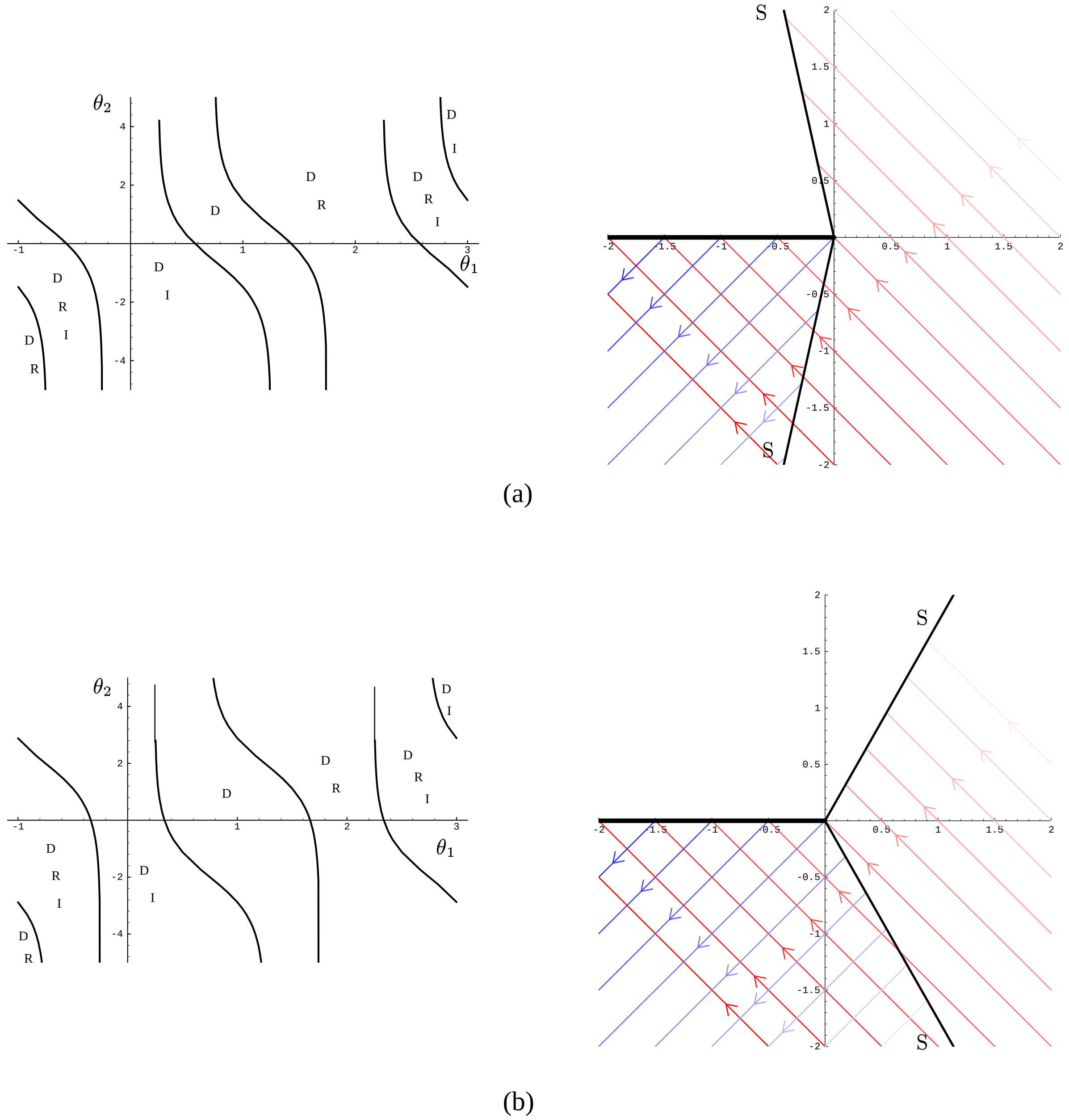


Figure 65: Stokes lines for the diffraction of a complex plane wave by a half-line with (a) $\alpha = \pi/4 + 0.6i$ and (b) $\alpha = \pi/4 + 2i$. The arrows indicate the normals to the level curves of $\mathrm{Re}(u)$, which is the direction of energy flow. These are also the level curves of $\mathrm{Im}(u)$, and thus indicate lines of constant magnitude of the wavefield. The incident field is red, the reflected field is blue, and the diffracted field has been omitted for clarity. The fading intensity of the colour corresponds to the exponential decay of the field.

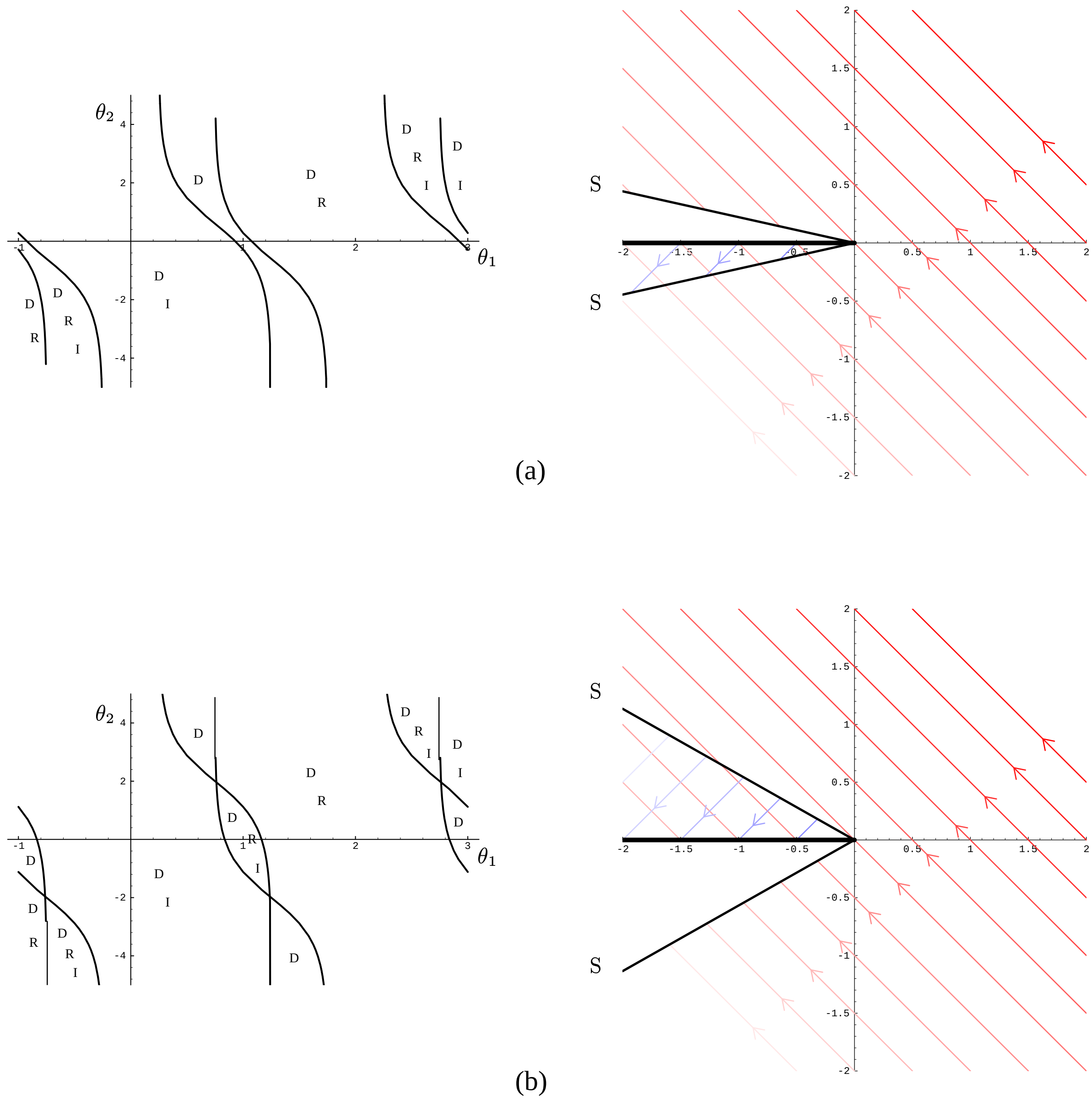
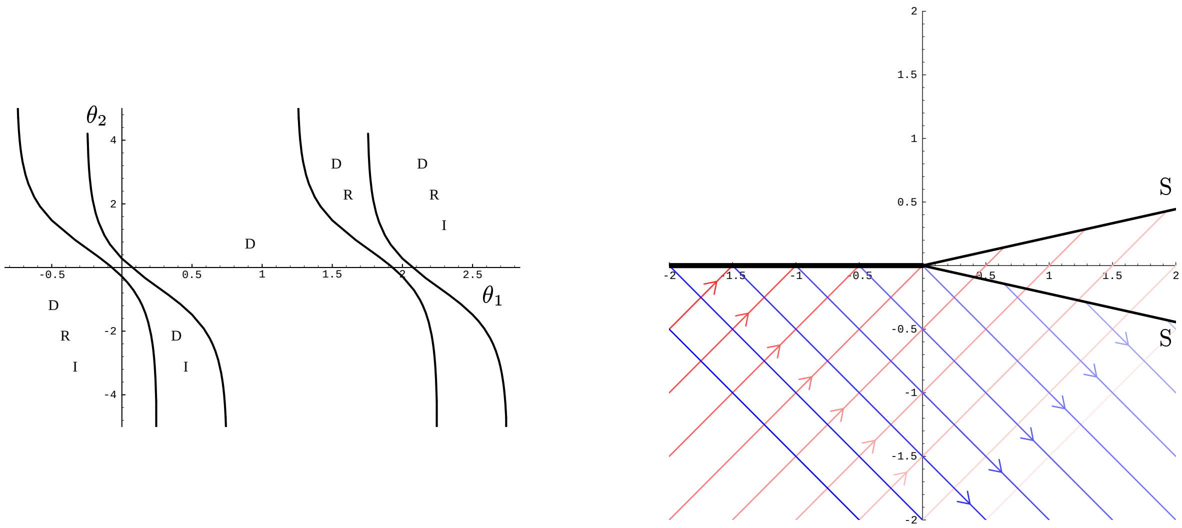
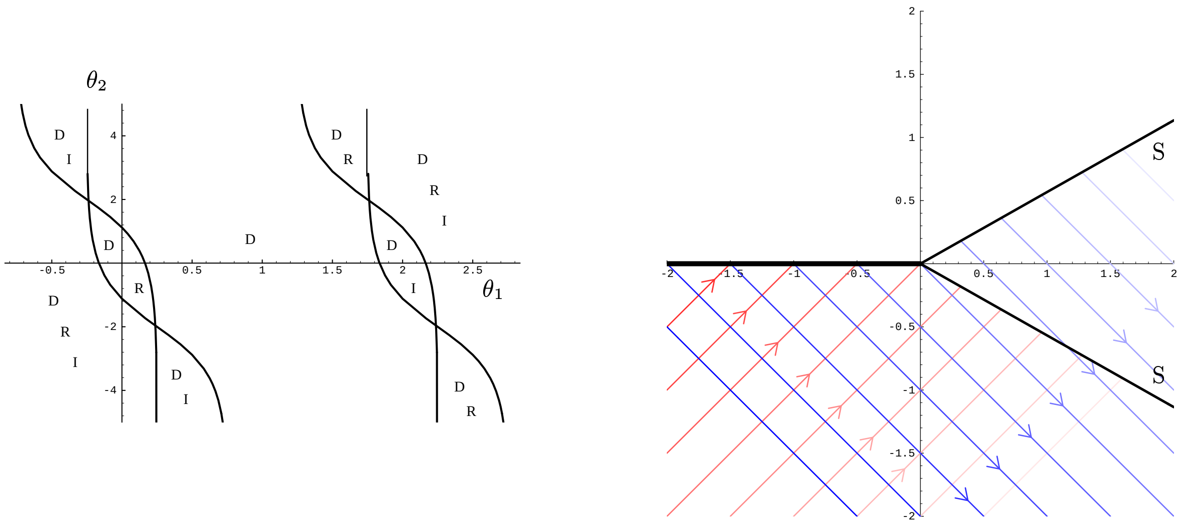


Figure 66: Stokes lines for the diffraction of a complex plane wave by a half-line with (a) $\alpha = \pi/4 - 0.6i$ and (b) $\alpha = \pi/4 - 2i$. The arrows indicate the normals to the level curves of $\mathrm{Re}(u)$, which is the direction of energy flow. These are also the level curves of $\mathrm{Im}(u)$, and thus indicate lines of constant magnitude of the wavefield. The incident field is red, the reflected field is blue, and the diffracted field has been omitted for clarity. The fading intensity of the colour corresponds to the exponential decay of the field.

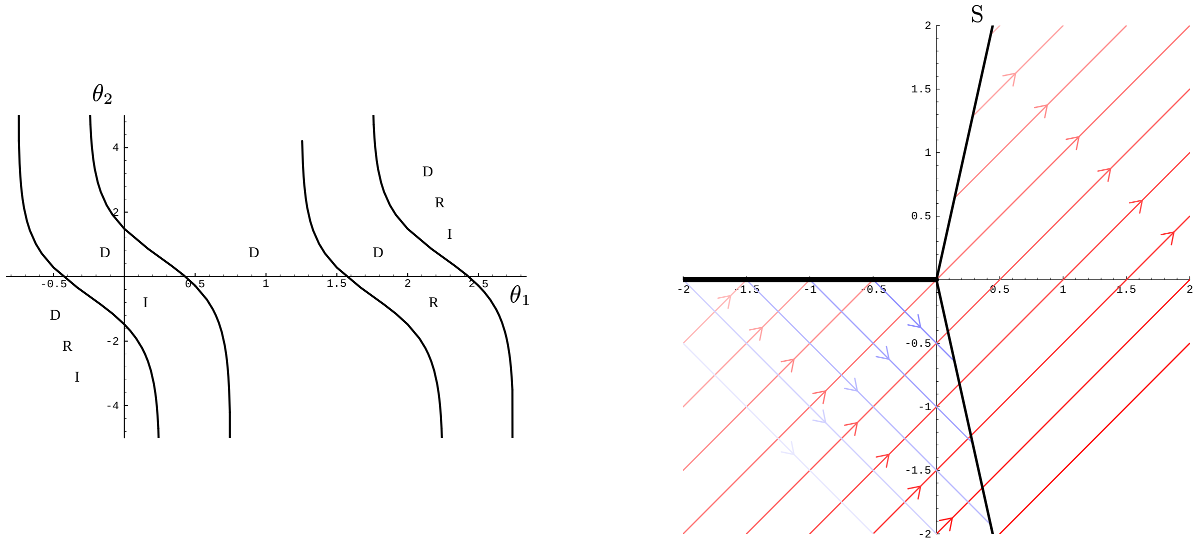


(a)

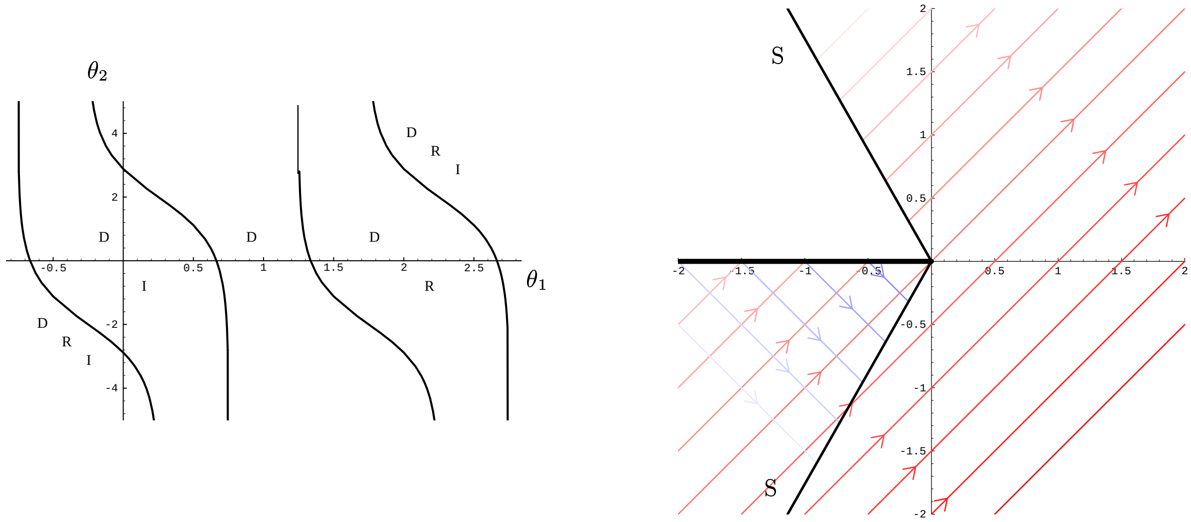


(b)

Figure 67: Stokes lines for the diffraction of a complex plane wave by a half-line with (a) $\alpha = 3\pi/4 + 0.6i$ and (b) $\alpha = 3\pi/4 + 2i$. The arrows indicate the normals to the level curves of $\mathrm{Re}(u)$, which is the direction of energy flow. These are also the level curves of $\mathrm{Im}(u)$, and thus indicate lines of constant magnitude of the wavefield. The incident field is red, the reflected field is blue, and the diffracted field has been omitted for clarity. The fading intensity of the colour corresponds to the exponential decay of the field.



(a)



(b)

Figure 68: Stokes lines for the diffraction of a complex plane wave by a half-line with (a) $\alpha = 3\pi/4 - 0.6i$ and (b) $\alpha = 3\pi/4 - 2i$. The arrows indicate the normals to the level curves of $\mathrm{Re}(u)$, which is the direction of energy flow. These are also the level curves of $\mathrm{Im}(u)$, and thus indicate lines of constant magnitude of the wavefield. The incident field is red, the reflected field is blue, and the diffracted field has been omitted for clarity. The fading intensity of the colour corresponds to the exponential decay of the field.

match it with the ray solution. The result is the addition of an additional diffracted ray field originating from the singularity. The inner analysis gives the directivity of this source, that is, the amplitude as a function of θ. We employ the same methodology when dealing with complex rays.

Consider a source located at the complex point (x_0, y_0). We parametrise the boundary $x < 0$ by

$$x = s, \qquad y = 0.$$

As in §2.3, the incoming field is given by

$$\begin{aligned}
\phi_{\text{inc}} &= H_0^{(1)}(k((x-x_0)^2+(y-y_0)^2)^{1/2}), &(332)\\
&\sim \frac{e^{ik((x-x_0)^2+(y-y_0)^2)^{1/2}}}{\sqrt{k}((x-x_0)^2+(y-y_0)^2)^{1/4}}, &(333)
\end{aligned}$$

which has the ray description

$$\begin{aligned}
x &= s + \frac{t(s-x_0)}{\sqrt{(s-x_0)^2+y_0^2}}, &(334)\\
y &= -\frac{ty_0}{\sqrt{(s-x_0)^2+y_0^2}}, &(335)\\
u_i &= \sqrt{(s-x_0)^2+y_0^2}+t, &(336)\\
A &= \frac{1}{(\sqrt{(s-x_0)^2+y_0^2}+t)^{1/2}}. &(337)
\end{aligned}$$

The reflected field has the ray description

$$\begin{aligned}
x &= s + \frac{t(s-x_0)}{\sqrt{(s-x_0)^2+y_0^2}}, &(338)\\
y &= \frac{ty_0}{\sqrt{(s-x_0)^2+y_0^2}}, &(339)\\
u_r &= \sqrt{(s-x_0)^2+y_0^2}+t, &(340)\\
A &= -\frac{1}{(\sqrt{(s-x_0)^2+y_0^2}+t)^{1/2}}, &(341)
\end{aligned}$$

which can be inverted to give

$$\phi_{\text{ref}} \sim \frac{e^{ik((x-x_0)^2+(y+y_0)^2)^{1/2}}}{\sqrt{k}((x-x_0)^2+(y+y_0)^2)^{1/4}}, \tag{342}$$

as expected. The ray that hits the tip of the half-line is given by

$$s = 0, \qquad x = -\frac{tx_0}{\sqrt{x_0^2+y_0^2}}, \qquad y = -\frac{ty_0}{\sqrt{x_0^2+y_0^2}},$$

so that the (complex) angle of incidence is given by

$$\alpha = \tan^{-1}(y_0/x_0). \tag{343}$$

To consider the inner diffraction problem we re-scale about the point $x = y = 0$ by setting $x = X/k$, $y = Y/k$ to give

$$\nabla^2\phi + \phi = 0. \tag{344}$$

The ray meeting the tip gives an incoming plane wave with angle of incidence α. Then, from (305), the inner solution is

$$\phi = \frac{e^{ik(x_0^2+y_0^2)^{1/2}}}{(x_0^2+y_0^2)^{1/4}} e^{-i\pi/4+ir}\,\mathrm{Fr}\left(-\sqrt{2r}\cos\left(\frac{\theta-\alpha}{2}\right)\right) - e^{-i\pi/4+ir}\,\mathrm{Fr}\left(-\sqrt{2r}\cos\left(\frac{\theta+\alpha}{2}\right)\right), \tag{345}$$

and the diffracted field to be included in the ray solution is

$$\frac{e^{ik(x_0^2+y_0^2)^{1/2}}}{(x_0^2+y_0^2)^{1/4}}\frac{e^{ikr-i\pi/4}}{2\sqrt{2\pi kr}}\left(\frac{1}{\cos((\theta+\alpha)/2)} - \frac{1}{\cos((\theta-\alpha)/2)}\right). \tag{346}$$

Having obtained our three rayfields, we now need to determine the Stokes surfaces in order to determine the shadow boundaries. We find that the phases of the diffracted and incident rays are equal along the curve

$$\mathrm{Re}\left[\left((x-x_0)^2+(y-y_0)^2\right)^{1/2}\right] = \mathrm{Re}\left[\left(x_0^2+y_0^2\right)^{1/2} + \left(x^2+y^2\right)^{1/2}\right] \tag{347}$$

while the phases of the diffracted and reflected rays are equal along the curve

$$\mathrm{Re}\left[\left((x-x_0)^2+(y+y_0)^2\right)^{1/2}\right] = \mathrm{Re}\left[\left(x_0^2+y_0^2\right)^{1/2} + \left(x^2+y^2\right)^{1/2}\right]. \tag{348}$$

The Stokes surfaces are the equal-phase surfaces on which the diffracted field is dominant. Thus the Stokes surfaces in real space, which were simply the shadow boundaries $\theta = \pm(\pi-\alpha)$ in the classical problem, and were rotated as in (330)-(331) for complex plane waves, are now both rotated and curved, as shown in an example in Fig. 69.

We note that the problem under consideration has an exact solution, obtained by formally complexifying the source point in the exact solution for a real source on a half-line [15]. We have used the problem merely to illustrate the procedures involved in the complex Geometrical Theory of Diffraction. As in the classical problem, it can be shown by explicit calculation that the ray solution obtained agrees with the ray expansion of the exact solution.

8 Conclusion

Motivated by the need for ever-more-accurate representations of high-frequency wavefields, we have tried to give a systematic description of a theory of complex rays in the frequency domain. Our aim has been to assemble a theory that is consistent in an applied mathematical sense and whose implementation is feasible in practical situations.

Geometrically, our starting point has been the interpretation of complex rays as manifolds in the complexified space of the independent variables in the Helmholtz equation. Usually the phase function is defined on a multiple-sheeted Riemann manifold in this complexified space, and the asymptotic analysis of successively higher terms in the ray expansion inevitably involves Stokes' phenomenon and the switching on and off of complex rayfields across suitably defined "Stokes surfaces". In this article, these are three-dimensional manifolds in a four-dimensional complex space. Although it is easy to find all the possible components of these surfaces by using the recipe (98), not all these "equal-phase" components are active in any particular problem. We have only been able to give a comprehensive delineation of the active components for some of the problems discussed in §5–7, including ones for which checks using explicit or Fourier-transform representations are available, but we are confident that most of the "active" structure for general smooth convex scattering problems can be found using (98) and (110).

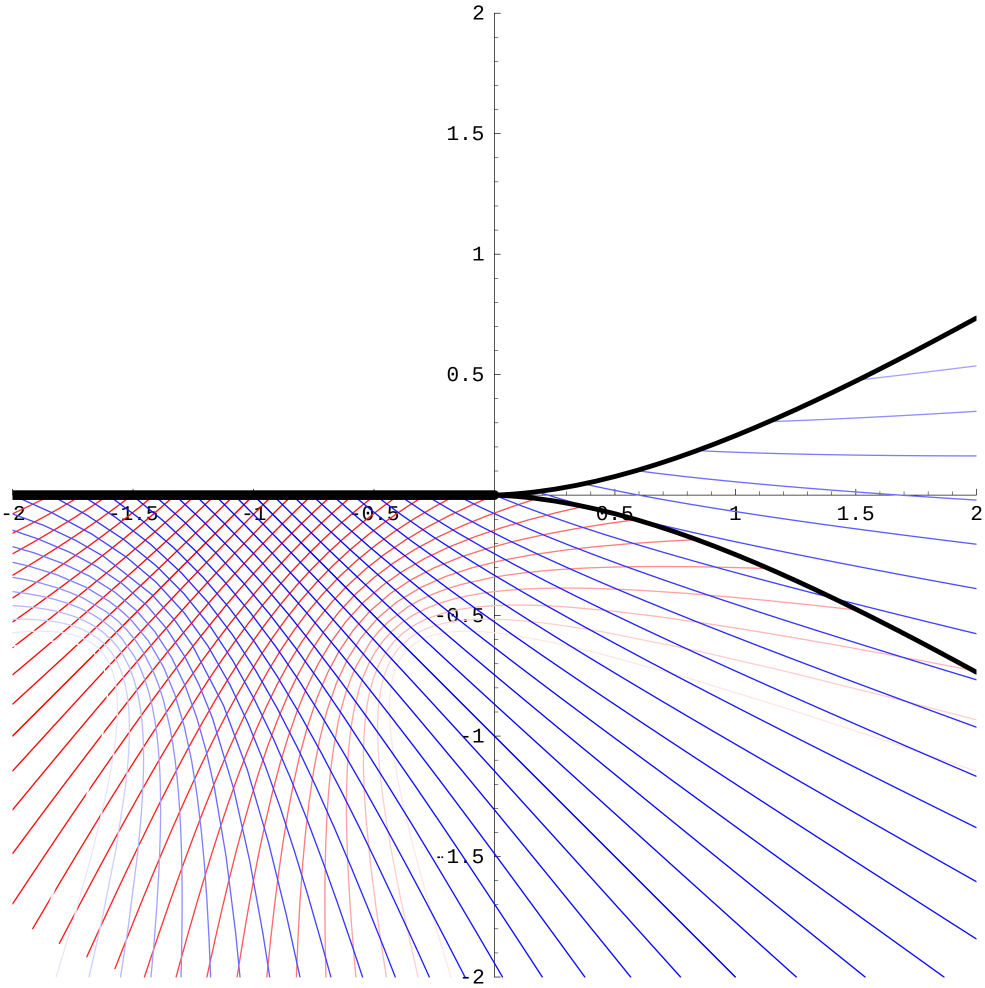


Figure 69: Stokes lines for the diffraction of a Gaussian beam by a half-line. With the notation of §2.3 for the beam $b_1 = 0$, $b_2 = 1$, and $\alpha = -\pi/4$, all with respect to an origin fixed at $(-1, 0)$. The lines are the normals to the level curves of $\mathrm{Re}(u)$, which is the direction of energy flow. These are also the level curves of $\mathrm{Im}(u)$, and thus indicate lines of constant magnitude of the wavefield. The incident beam is shown in red, the reflected beam in blue, and the diffracted rays have been omitted for clarity. The fading intensity of the colours corresponds to the exponential decay of the wavefield.

Although our work relies heavily on analytic continuation we have found that it can be applied even in the presence of nonanalytic boundaries, for which the Geometrical Theory of Diffraction is the appropriate tool in real ray theory. We find that the Geometrical Theory of Diffraction can still be used, but the singularities which are generically present outside physical space due to the process of analytic continuation are now also to be found at the points of nonanalyticity of the boundary. Nonetheless, quite surprising complex ray pictures emerge from such singularities, for example as illustrated in §7 for half-lines illuminated by plane waves and Gaussian beams.

In summary we envisage the following general procedure for tackling problems involving complex rays:

1. Solve the eikonal equation to identify all the "outer" incident, reflected and transmitted rays, real and complex.

2. Identify singularities in the rayfields such as tangency points (real or complex) and nonanalytic boundary points.

3. Solve the inner diffraction problems near these points to determine the diffracted/creeping fields which must be added to the outer ray expansion[7].

4. Identify the equal phase surfaces between pairs of rayfields.

5. Identify which of these equal phase surfaces are Stokes surfaces.

The main difficulty arises in (5) above, for whose implementation we have three tools at our disposal:

1. *Either* some of the local diffraction problems may have explicit integral solutions available whose Stokes surface structure can be determined by inspection. Whether or not branches of the equal-phase surfaces are Stokes surfaces in the ray solution can then be identified by using continuity arguments;

2. *or* in some regions of space we may know which rays we expect to be present on physical grounds, and we can then use consistency arguments to determine which equal phase surfaces are Stokes surfaces;

3. *or* it may be possible to use information from the tail of the asymptotic expansion of one rayfield to determine whether it is responsible for the switching on of another ray through Stokes phenomenon.

In practical terms (1) and (2) above are more useful than (3), although there may be some occasions in which (3) proves useful. Note in particular that (3) implies that a ray contribution which has a converging asymptotic expansion (including in particular a terminating asymptotic expansions such as a plane wave) can never be responsible for switching on or off another wavefield.

Apart from the above-mentioned incompleteness of our characterisation of which equal-phase surfaces are in fact "active" Stokes surfaces, other problems that await attention include the illumination of boundaries where the curvature changes sign, problems for the Helmholtz equation in interior domains, problems in three space dimensions or in the time domain, and the generalisation of complex ray theory to more complicated partial differential equations.

[7]Note that as we saw in §6.3, the inner solution near a complex tangency point is simply the same as the solution near a real tangency point after a suitable complex rotation of coordinates, so that the creeping rays generated by complex tangencies will be the same as those generated by real tangencies.

Our analyses in §6.2, §6.4 and §7 show clearly that while incident, reflected, creeping and head rays at the boundary in real space are all real, complex rays can be excited at points of the complexified boundary that do intersect real space. Moreover, by using (98) and (110) it is possible to identify the curves (Stokes surfaces) in real space across which the complexified wavefields are switched on and off.

One methodological aspect of our work that deserves mention concerns the dependence of our prediction on our particular choice of (1) as the ansatz for ϕ. If, for example, we had written $u = iu'$ in (1), the restriction of the rays associated with u' to real values of s and t would have been different from that for u. However, if all rays are considered as complex rays, they are the same for u' as for u: they simply have different parameterisations. Thus while the choice of ansatz is at the discretion of the investigator, all predictions about ϕ, in particular its Stokes-surface structure, are ansatz-independent. Our choice in this paper has been that which accords with most of ray theory literature.

The question of the physical significance of complex rays remains a contentious one. In an experimentally measured wavefield, or in a numerical solution of Helmholtz equation, the Stokes lines will rarely be obvious: they will usually only emerge if the optimally truncated ray expansion is subtracted from the measured or calculated value, and then only if this value is sufficiently accurate. Of course, as a caustic is approached the complex ray contributions become relatively larger; thus, if a Stokes surface intersects real space nearby, as occurs say in the case of a cusped caustic [44] and as occurred in §6.2 near $x = x_0$, then the switching phenomenon becomes easier to discern.

Finally we note that in cases such as the insonification of a cylinder by a complex plane wave in §6.4, while it may not be too important to know exactly *where* the Stokes switching has occurred, it is crucial to know that it *has* occurred, since the wavefield switched on will grow to dominate the switching wavefield at other points in space.

Acknowledgements. We would like to express gratitude to our students Dr B. Smith and Dr V. Saward for all their work in support of this research and to Dr Colin Sillence of British Aerospace who motivated our early investigations. Many other people have made all kinds of helpful comments during the preparation of the paper, especially Prof. V. M. Babic, Prof. Sir M. V. Berry, Dr C. J. Chapman, Dr A. P. Kiselev, Prof. J. R. King, and not least the late Prof. G. R. Wickham. Finally, we would like to express especial gratitude to the Smith Institute, Guildford, for providing facilities and resources which enabled much of the early work in this paper to be carried out.

References

[1] V. M. Babič and V. S. Buldyrev. *Short-Wavelength Diffraction Theory — Asymptotic Methods*, volume 4 of *Springer Series on Wave Phenomena*. Springer-Verlag, 1991.

[2] V. M. Babič and N. Y. Kirpičnikova. *The Boundary Layer Method in Diffraction Problems*. Springer-Verlag, Berlin, 1979.

[3] B. B. Baker and E. T. Copson. *The Mathematical Theory of Huygens' Principle*. Clarendon Press, Oxford, 2nd edition, 1950.

[4] M. V. Berry. Uniform asymptotic smoothing of Stokes discontinuities. *Proc. Roy. Soc. Lond. A*, 422:7–21, 1989.

[5] M. V. Berry and C. J. Howls. Hyperasymptotics. *Proc. Roy. Soc. Lond. A*, 430:653–667, 1990.

[6] H. L. Bertoni, A. C. Green, and L. B. Felsen. Shadowing an inhomogeneous plane wave by an edge. *J. Opt. Soc. Am.*, 68(7):983–989, 1978.

[7] N. Bleistein. *Mathematical Methods for Wave Phenomena*. Academic Press, 1984.

[8] R. N. Buchal and J. B. Keller. Boundary layer problems in diffraction theory. *Comm. Pure Appl. Math.*, 13:85–114, 1960.

[9] C. J. Chapman. Sound radiation from a cylindrical duct, part 1: Ray structure of the duct modes and of the external field. *J. Fluid Mech.*, 281:293–311, 1994.

[10] A. B. O. Daalhuis and F. W. J. Olver. Hyperasymptotic solutions of second-order linear differential equations. *Methods and Applications of Analysis*, 2:173–197, 1995.

[11] G. A. Deschamps. Gaussian beam as a bundle of complex rays. *Electron. Lett.*, 7(23):684–685, 1971.

[12] R. B. Dingle. *Asymptotic Expansions: their Derivation and Interpretation*. Academic Press, 1973.

[13] P. D. Einziger and L. B. Felsen. Evanescent waves and complex rays. *IEEE Trans. Antennas Prop.*, 30(4):594–605, 1982.

[14] L. B. Felsen. Evanescent waves. *J. Opt. Soc. Am.*, 66:751–760, 1976.

[15] A. C. Green, H. L. Bertoni, and L. B. Felsen. Properties of the shadow cast by a half-screen when illuminated by a Gaussian beam. *J. Opt. Soc. Am.*, 69(11):1503–1508, 1979.

[16] E. Heyman and L. B. Felsen. Evanescent waves and complex rays for modal propagation in curved open waveguides. *SIAM J. Appl. Math.*, 43(4):855–884, 1983.

[17] D. S. Jones. *The Theory of Electromagnetism*. Pergamon Press, 1964.

[18] D. S. Jones. The Kontorovich-Lebedev transform. *J. Inst. Math. Appl.*, 26:133–141, 1980.

[19] J. B. Keller. Diffraction by a convex cylinder. In *URSI Michigan Electromagnetic Wave Theory Symposium*, pages 312–321, 1955.

[20] J. B. Keller. A geometrical theory of diffraction. In *Proc. Symposia Appl. Math.*, volume 8, pages 27–32. McGraw-Hill, 1958.

[21] J. B. Keller. The geometrical theory of diffraction. *J. Opt. Soc. Am.*, 52:116–130, 1962.

[22] J. B. Keller. Semiclassical mechanics. *SIAM Review*, 27(4):485–504, 1985.

[23] J. B. Keller and F. C. Karal, Jr. Surface wave excitation and propagation. *J. Appl. Phys.*, 31(6):1039–1046, 1960.

[24] J. B. Keller and R. M. Lewis. *Asymptotic Methods for Partial Differential Equations: The Reduced Wave Equation and Maxwell's Equations*, volume 1 of *Surveys in Applied Mathematics*, pages 1–82. Plenum Press, New York, 1995.

[25] J. B. Keller and W. Streifer. Complex rays with an application to Gaussian beams. *J. Opt. Soc. Am.*, 61:40–43, 1971.

[26] Y. A. Kravtsov. Complex rays and complex caustics. *Radiophys. Quantum Electron.*, 10:719–730, 1971.

[27] Y. A. Kravtsov and Y. I. Orlov. *Geometrical Optics of Inhomogeneous Media*, volume 6 of *Springer Series on Wave Phenomena*. Springer Verlag, 1990.

[28] M. Leontovich and V. Fock. Solution of the problem of the proagation of electromagnetic waves along the Earth's surface by the method of the parabolic equation. *J. Phys.*, 10:13–24, 1946.

[29] B. R. Levy and J. B. Keller. Diffraction by a smooth object. *Comm. Pure Appl. Math.*, 12:159–209, 1959.

[30] R. M. Lewis, N. Bleistein, and D. Ludwig. Uniform asymptotic theory of creeping rays. *Comm. Pure Appl. Math.*, 20:295–328, 1967.

[31] D. Ludwig. Uniform asymptotic expansions at a caustic. *Comm. Pure Appl. Math.*, 19:215–250, 1966.

[32] J. B. McLeod. Smoothing of Stokes discontinuities. *Proc. Roy. Soc. Lond. A*, 437:343–354, 1992.

[33] R. E. Meyer. A simple explanation of the Stokes phenomenon. *SIAM Review*, 31(3):435–445, 1989.

[34] A. B. Olde Daalhuis, S. J. Chapman, J. R. King, J. R. Ockendon, and R. H. Tew. Stokes phenomenon and matched asymptotic expansions. *SIAM J. Appl. Math.*, 55(6):1469–1483, 1995.

[35] F. W. J. Olver. *Asymptotics and Special Functions*. Academic Press, New York, 1974. Reprinted in AKP Classics, 1997.

[36] R. B. Paris and A. D. Wood. Stokes phenomenon demystified. *IMA Bulletin*, 31:21–28, 1995.

[37] H. Segur, S. Tanveer, and H. Levine, editors. *Asymptotics Beyond All Orders*, volume 284 of *NATO ASI Series B*. Plenum Press, 1991.

[38] B. J. Smith. *A Complex Ray Approach to the Acoustics of Fluid-Loaded Structures*. PhD thesis, University of Nottingham, 1995.

[39] B. J. Smith and R. H. Tew. Surface wave excitation on a flexible structure. *Wave Motion*, 21(3):277–289, 1995.

[40] A. Sommerfeld. Mathematische Theorie der Diffraction. *Mathematische Annalen*, 47:317–374, 1896.

[41] R. H. Tew. Non-specular reflection phenomena at a fluid-solid boundary. *Proc. Roy. Soc. Lond. A*, 437(1900):433–449, 1992.

[42] R. H. Tew and J. R. Ockendon. Multiple scale expansions near critical rays. *SIAM J. Appl. Math.*, 52:327–336, 1992.

[43] W.-Y. D. Wang and G. A. Deschamps. Application of complex ray tracing to scattering problems. *Proc. IEEE*, 62(11):1541–1551, 1974.

[44] F. J. Wright. The Stokes set of the cusp diffraction catastrophe. *J. Phys.*, 13:2913–2928, 1980.

[45] E. Zauderer. Wave propagation around a convex cylinder. *J. Math. Mech.*, 13:171–186, 1964.

[46] E. Zauderer. Boundary layer and uniform asymptotic expansions for diffraction problems. *SIAM J. Appl. Math.*, 19:575–600, 1970.