IOP PUBLISHING EUROPEAN JOURNAL OF PHYSICS

Eur. J. Phys. **34** (2013) 1423–1437 doi:10.1088/0143-0807/34/6/1423

Raman and the mirage revisited: confusions and a rediscovery

M V Berry

H H Wills Physics Laboratory, Tyndall Avenue, Bristol BS8 1TL, UK

E-mail: asymptotico@bristol.ac.uk

Received 14 April 2013, in final form 7 August 2013
Published 16 September 2013
Online at stacks.iop.org/EJP/34/1423

Abstract

Raman argued that in a continuously varying layered medium, such as air above a hot road, a ray that bends so as to become horizontal must remain so, implying that the reflection familiar in the mirage cannot be explained by geometrical optics. This is a mistake, as standard ray curvature arguments demonstrate. But a simple limiting process, in which the smoothly varying refractive index is approximated by a stack of thin discrete layers, is not quite straightforward because it involves a curious singularity, related to the level ray envisaged by Raman. In contrast to individual rays, families of rays possess caustic (focal) singularities. These can be calculated explicitly for two families of rays that are relevant to the mirage. Only exceptionally does the locus of reflection (lowest points on the rays) coincide with the caustics. Caustics correspond to the 'vanishing line', representing the limiting height of objects that can be seen by reflection. For these two families, the waves that decorating mirage caustics are described by the universal Airy function, and can be calculated exactly.

1. Introduction

The confusions of great scientists can be instructive. A fascinating example is two papers written in 1959 by Raman with his nephew Pancharatnam [1], and by Raman alone [2], concerning the mirage images of distant objects commonly seen reflected in a road on a hot day. As Raman declares: 'The theory of the mirage which is usually accepted purports to base itself on geometrical optics'. The explanation that is 'usually accepted'—and which is correct as we will see—is that the temperature decrease above the heated road is associated with a refractive index of air that increases with height, and this index gradient bends light rays so they are concave upwards: refraction leads to the appearance of reflection.

Raman disagreed, regarding the geometrical-optics explanation as 'inadequate and unsatisfactory' and 'a kind of make-believe'. The basis of this opinion was that, in a refractive-index gradient, 'a pencil of rays travelling obliquely through the stratified medium, would, according to Snel's[1] law of refraction, be progressively deviated until it reaches a layer at

[1] 'Snel' is the correct spelling of the scientist's full name: Willebrord Snel van Royen. The more common 'Snell' is probably a mistaken anglicization originating in the Latin version Snellius: see e.g. David Park's authoritative history of optics: *The Fire within the Eye* (1997 Princeton University Press), where the spelling is Snel throughout.

which its course becomes tangential to the plane of the stratifications; thereafter, it would continue on a course parallel to the stratifications. No question of total reflection can therefore arise'.

This question of 'the path of a level ray' had been considered many years earlier. Writing in 1873, James Thomson (elder brother of Lord Kelvin) [3] addressed a question he attributed to Purser in 1863: 'To find whether a ray of light passing infinitely nearly horizontally through the atmosphere will be bent with a finite curvature, or not bent at all'. Thomson gave an argument based on regarding the refractive-index gradient as the limit of a stack of discrete layers, with the index constant in each, and arrived at the correct conclusion that the ray will be curved. But Thomson is not quite convinced: '... there is something perplexing, or not quite satisfactory to the mind, in taking this final step to the perfectly level ray; for as soon as the inclination of the ray becomes zero the whole foundation and framework of the investigation fails, there, being then no oblique passage of a ray from one lamina into another ...'. Therefore he gave an additional argument based on velocity gradients and wavefronts; a similar explanation had been given in 1856 by Bravais [4]. And writing in the same year, Everett [5], after again giving a correct analysis, uncompromisingly referred to 'The mistake of supposing that a ray can pursue a straight course parallel to planes of equal index in a continuously varying medium. The contrary was pointed out so long ago as 1799 and 1800 by Vince [6] and Wollaston [7] in the Philosophical Transactions, but appears to have since dropped out of mind' (citations added).

The standard argument, that a continuous refractive-index gradient curves a ray in a way that is analogous to the way a force bends the trajectory of a massive particle, is reviewed in section 2. It might seem that there is nothing that more to be said about this problem which was solved so long ago. But the limit of ray propagation through a stack of discrete constant-index layers is not quite straightforward, and involves a curious singularity that evokes the level ray envisaged by Raman; this is examined in detail in section 3.

The erroneous opinion that geometrical optics is inadequate led Raman and Pancharatnam [1] to analyse the mirage using wave optics, and to the observation that '... the situation in the vicinity of the limiting layer would resemble that well known in physical optics to appear in the vicinity of caustic surfaces ...'. Their analysis was correct, and indeed the importance of caustics associated with the mirage had been emphasized as long ago as 1810, in a comprehensive study by Biot [8]. Nevertheless, their statement is also misleading: although the mirage does involve caustics, it is only in exceptional cases that these are associated with the 'limiting layer' where reflection occurs. This is immediately obvious from the observation that caustics are commonly formed not only in varying media but also in a uniform medium, from rays that are not curved at all. Two familiar examples are: the caustic formed by the straight rays reflected from the inner surface of a coffee cup, where the cusped caustic is located away from the surface where the reflection occurs; and the rainbow, in which the caustic extends into the far field, where it can intersect the viewer's eye kilometres from the ray reflection inside the raindrop. In section 4, mirage caustics are examined within geometrical optics, as focal surfaces formed by families of rays, in two cases: one where the caustic coincides with the locus of ray reflections, and one where it does not.

There is a curious postscript to this story. The wave treatment by Raman and Pancharatnam, and the association with caustics for the special case they examine, are both correct. But they seemed unaware that their analysis reproduces exactly the wave that was discovered by Airy [9] in 1838, and which decorates any caustic, not just those associated with mirages. This rediscovery of the Airy function is explained in section 5, in which the waves corresponding to the two ray families considered in section 4 are derived exactly (with some mathematical details given in the appendix). In one of these (waves from a point source in an uniform

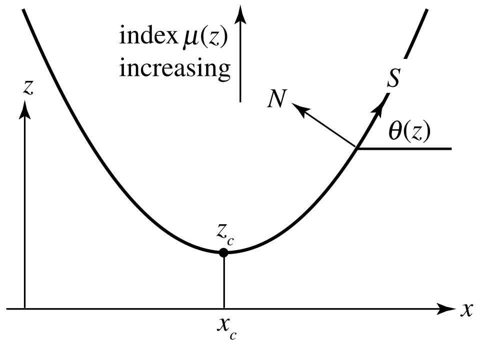

Figure 1. Geometry and notation for continuously bending ray.

gradient of squared refractive index), the exact formula is known in one space dimension (though not as well as it deserves to be) but seems unfamiliar in three.

In the concluding section 6, I suggest several ways in which this study might be helpful in teaching optics, or physics more generally.

My focus in this paper is to correct misconceptions in the basic optical theory of mirages, not to describe the many associated multiple-image phenomena. These are well discussed elsewhere [10–14]. A very useful bibliography on all aspects of mirages has been compiled by A T Young [15, 16].

2. Review of ray reflection in a smoothly varying medium

In the flat earth approximation, which suffices for present purposes, the refractive index above a hot road can be modelled by a smooth function $\mu(z)$ of height z that initially increases and then reaches a constant value. For a ray whose direction at height z makes an angle $\theta(z)$ with the horizontal (figure 1) (note: not the vertical as is usual), Snel's law gives the conserved quantity

$$\mu(z)\cos(\theta(z)) = \mu(z_c) \equiv \mu_c \quad (\theta(z_c) = 0), \tag{2.1}$$

in which z_c is the height at which the ray is horizontal, that is, the reflection level. This represents the conservation of horizontal photon momentum for a medium invariant under x translation.

Differentiation gives

$$\partial_z\mu(z)\cos(\theta(z)) = \mu(z)\sin(\theta(z))\partial_z\theta(z), \tag{2.2}$$

and hence, after changing variables from z to x via

$$\partial_x z(x) = \tan[\theta(z(x))], \tag{2.3}$$

the acceleration equation

$$\partial_x^2 z = \frac{1}{\cos^2\theta}\partial_z\theta\,\partial_x z = \frac{\partial_z \log\mu}{\cos^2\theta} = \frac{\partial_z\mu^2}{2\mu_c^2}. \tag{2.4}$$

A further transformation, to ray arc length S and distance N normal to the ray (figure 1) using

$$\mathrm{d}z = \mathrm{d}S\sin\theta = \mathrm{d}N\cos\theta \tag{2.5}$$

gives, finally

$$\text{ray curvature} = \partial_S \theta = \partial_N \log \mu = \text{logarithmic normal derivative of refractive index}. \tag{2.6}$$

This is the standard equation determining the path of a ray in geometrical optics [17], and in the form (2.6) it is valid in any smoothly varying medium, not only layered ones: the index can be a general function of position. The result goes back at least to Wollaston in 1800 [7], who understood that it excludes the persistently horizontal ray envisaged by Raman: 'If the density of any medium varies by parallel indefinitely thin strata, any rays of light moving through it *in the direction of the strata*, will be made to deviate during their passage, and their deviations will be in proportion to the increments of density where they pass. For each ray will be bent towards the denser strata, by a refracting force proportioned to the difference of the densities above and below the line of its passage ... '.

For an explicit form of the rays, we need not solve (2.4) or (2.6), but can use the conservation law (2.1) directly. Thus

$$\partial_x z = \pm\sqrt{\sec^2\theta - 1}, \text{ i.e. } \mathrm{d}x = \pm \frac{\mathrm{d}z}{\sqrt{(\mu(z)/\mu_c)^2 - 1}}, \tag{2.7}$$

leading, after integration, to

$$x = x_c \pm \int_{z_c}^{z} \frac{\mathrm{d}z'}{\sqrt{(\mu(z')/\mu_c)^2 - 1}}. \tag{2.8}$$

Where the ray is nearly horizontal, we can expand the refractive index about the reflection level, that is

$$\mu(z) \approx \mu_c + (z - z_c)\partial_z \mu(z_c) \equiv \mu_c + (z - z_c)\mu_c', \tag{2.9}$$

to get the locally parabolic ray track

$$z(x) \approx z_c + \frac{\mu_c'}{2\mu_c}(x - x_c)^2. \tag{2.10}$$

If we model μ^2 by a function that is linear everywhere (an unrealistic model for the mirage!), that is,

$$\mu^2(z) = Az, \tag{2.11}$$

the ray is parabolic everywhere, not just near the minimum:

$$z(x) \approx z_c + \frac{1}{4z_c}(x - x_c)^2. \tag{2.12}$$

This idealized mirage ray is then precisely analogous the upside-down trajectory of a particle projected horizontally and falling under gravity, a point that has often been made (e.g. as Wollaston's 'refracting force', by Laplace (p277ff of [18]), and more recently in [19]). The analogy was familiar to Newton, who in Principia describes '... the trajectories of bodies which are extremely like the trajectories of rays'. On the analogy, the mistakenly envisaged persistently horizontal ray would correspond to an obliquely thrown ball reaching its maximum height and then continuing to travel horizontally instead of falling, violating reversibility as well as experience.

3. Rays in a discretely layered medium: seeking the level ray

We now try to reproduce the curved rays in a smoothly varying medium by caricaturing the refractive index profile as a stack of horizontal layers of thickness Δ, with the index constant in

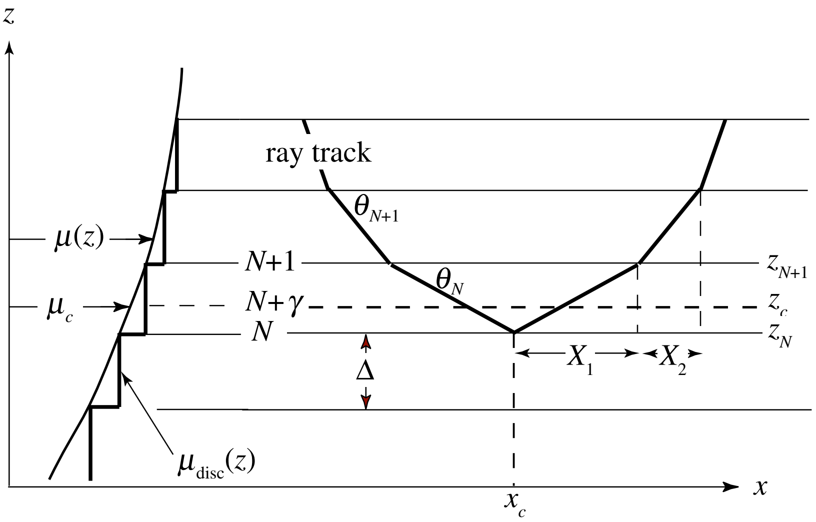

Figure 2. Discretely sampled refractive index $\mu_{\rm disc}(z)$ (equation (3.2)), and geometry and notation for ray refracted and totally internally reflected by the corresponding discrete layers n, with $n = N$ being the critical layer, at which the ray is reflected.

each, and examining the limiting process $\Delta \to 0$, using one of several possible discretizations. Referring to figure 2, the layers are bounded by z_{n-1} and z_n where

$$z_n = n\Delta, \tag{3.1}$$

and the discretized index is

$$\mu_{\rm disc}(z) = \mu\left(\Delta \mathrm{int}\left(\frac{z}{\Delta}+1\right)\right),$$
$$\text{i.e. } \mu_{\rm disc}(z) = \mu(z_n) \equiv \mu_n \quad \text{for} \quad z_{n-1} < z \leqslant z_n, \tag{3.2}$$

in which int(x) denotes the integer part of x. Snel's law (2.1) now becomes

$$\cos\theta_n = \frac{\mu_c}{\mu_{n+1}}, \tag{3.3}$$

with the index $n = N$ of the critical layer, below which the ray cannot penetrate, being defined by

$$\mu_N < \mu_c \leqslant \mu_{N+1}. \tag{3.4}$$

Rays are now a set of straight segments, each labelled by m with $m = 1$ denoting the lowest. The horizontal location of the end of the mth segment is (figure 2)

$$x_{N+m} = x_c + \sum_{l=1}^{m} X_l, \quad \text{where} \quad X_l = \Delta\cot\theta_{N+l-1} = \frac{\Delta}{\sqrt{(\mu_{N+l}/\mu_c)^2 - 1}}. \tag{3.5}$$

Again we expand the index (now discrete) about the critical value:

$$\mu_{N+l} \approx \mu_c + (z_{N+l} - z_c)\mu_c' = \mu_c\left(1 + \left(n_c + l - \frac{z_c}{\Delta}\right)\frac{\mu_c'\Delta}{\mu_c}\right). \tag{3.6}$$

To proceed further, it is necessary to specify the precise location of the critical index within the critical layer, as a fraction γ of the thickness:

$$\frac{z_c}{\Delta} = N + \gamma, \quad (0 < \gamma \leqslant 1). \tag{3.7}$$

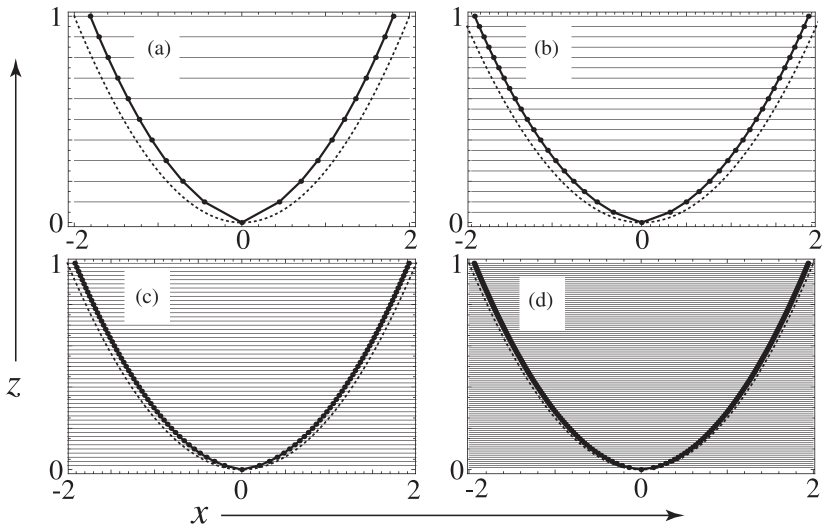

Figure 3. Convergence of discrete rays onto continuous-ray parabola (2.12) for $z_c = 1$ (dotted), for $\gamma = 0.5$ and (a) $\Delta = 0.1$, (b) $\Delta = 0.05$, (c) $\Delta = 0.02$, (d) $\Delta = 0.01$.

Thus, from (3.5) and (3.6), the location of the mth segment is

$$x_{N+m} = x_c + K \sum_{l=1}^{m} \frac{1}{\sqrt{l-\gamma}}, \quad \text{where} \quad K = \sqrt{\frac{\mu_c \Delta}{2\mu_c'}}. \tag{3.8}$$

Ignoring complications, the smallness of the constant K for thin layers allows the sum to be replaced by an integral, leading to

$$x(z) \approx x_c + 2K\sqrt{m} = x_c + 2K\sqrt{\frac{z - z_c}{\Delta}} = x_c + \sqrt{\frac{2\mu_c}{\mu_c'}(z - z_c)}, \tag{3.9}$$

and thus reproducing the curved ray (2.10) in typical cases. Figure 3 illustrates the condensation of the segmented trajectory onto the smooth locally parabolic one as $\Delta \to 0$.

But there is a complication—a sense in which the continuum limit is weakly singular. It occurs when γ is very close to unity, so that the ray in the last layer is almost horizontal—exactly the situation envisaged by Raman. In this case, the term $l = 1$ in (3.8) is not small, and must be separated from the terms $l = 2, 3, \ldots$ whose sum can still be replaced by an integral. Then the lowest ray can be arbitrarily long, and the continuum approximation fails. The subsequent terms $l = 2, 3, \ldots$ correspond to segments with slope of order $\sqrt{\Delta}$. If we define the lowest ray ($l = 1$) as near-horizontal if its slope is smaller than Δ, this corresponds to $1 - \gamma < \Delta$. These continuum-approximation-violating intervals in γ exist for all Δ, however small. But the intervals get smaller as $\Delta \to 0$. For any fixed γ, however small, the term $l = 1$ in (3.8) vanishes as $\Delta \to 0$, and the limiting ray becomes the local parabola of the continuum approximation. This is illustrated in figure 4.

A further point needs to be discussed, concerning reflections at the refractive index discontinuities between the layers. These reflections occur because the transfer of light intensity (here denoting $|\psi|^2$ for a wave ψ) across a discontinuity—that is, a change of refractive index over a distance small compared with the wavelength—is strictly outside the geometrical-optics approximation. An incident ray splits into two: a transmission and a reflection, with the separation determined by the Fresnel coefficients of wave physics. As Raman correctly points

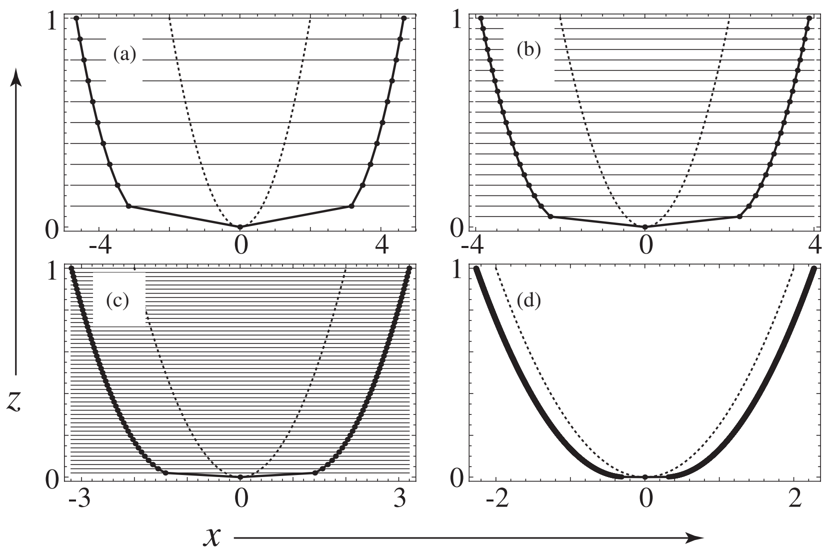

Figure 4. As figure 3, for $\gamma = 0.99$, showing slow convergence of near-horizontal ray onto continuous-ray parabola for (a) $\Delta = 0.1$, (b) $\Delta = 0.05$, (c) $\Delta = 0.02$, (d) $\Delta = 0.001$ (in (d), the layer lines are omitted for clarity). Note the different horizontal scales: all the parabolas are the same.

out, these reflections are small if the refractive-index difference between adjacent layers is small, except for the lowest ray which of course is totally reflected.

To understand this in a little more detail, it will suffice to use the scalar approximation, in which the intensity reflection coefficient at the interface z_n is [17]

$$R_n = \left|\frac{\sin\theta_{n-1} - \sin\theta_n}{\sin\theta_{n-1} + \sin\theta_n}\right|^2 = \left|\frac{\sqrt{(\mu_n/\mu_c)^2 - 1} - \sqrt{(\mu_{n+1}/\mu_c)^2 - 1}}{\sqrt{(\mu_n/\mu_c)^2 - 1} + \sqrt{(\mu_{n+1}/\mu_c)^2 - 1}}\right|^2, \quad (3.10)$$

where we have used Snel's law (3.3). Near the lowest layer $n = N$ this becomes, using (3.6) and (3.7),

$$R_{N+l} \approx \left|\frac{\sqrt{l-\gamma} - \sqrt{l+1-\gamma}}{\sqrt{l-\gamma} + \sqrt{l+1-\gamma}}\right|^2. \quad (3.11)$$

For $l = 0$, $\sqrt{l-\gamma}$ is imaginary, so $R_N = 1$: the ray is totally reflected as expected. For higher rays, the limiting form

$$R_{N+l} \underset{l\gg 1}{\to} \frac{1}{16l^2} = \frac{\Delta^2}{16(z-z_c)^2} \quad (3.12)$$

indicates that as $\Delta \to 0$ the reflections vanish for a fixed distance above z_c, however small. Therefore this admittedly peculiar discretization, with the rays treated geometrically by Snel's law and the reflections treated by wave theory, does lead to the combination of 'refraction leading to reflection' that underlies the mirage phenomenon.

4. Families of rays: caustics and reflection not the same, and the 'vanishing line'

So far, we have considered isolated rays. But optical fields correspond to *families* of rays, which have a holistic property not possessed by any individual ray: they typically form

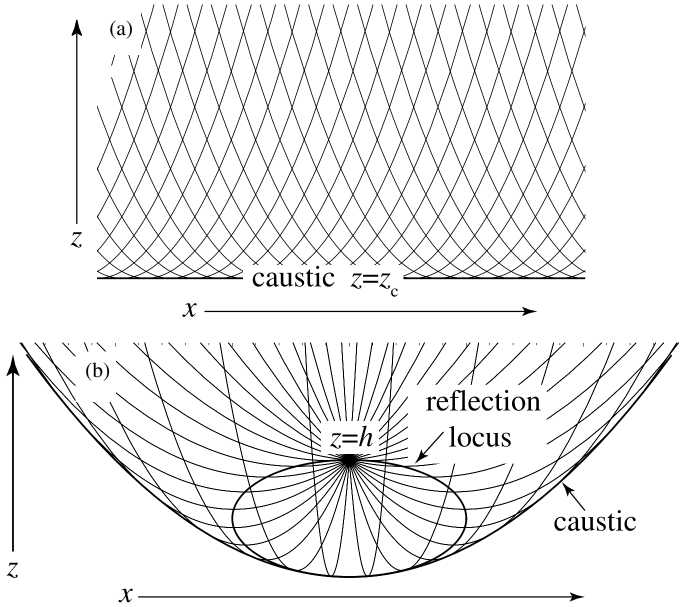

Figure 5. Families of mirage rays: (a) translationally invariant caustic formed by parallel rays; (b) paraboloidal caustic formed by rays from a point source, and locus of reflection points on the rays.

caustics, surfaces in space that are envelopes of the ray family, corresponding to focusing. As Stavroudis [20] eloquently puts it: 'The caustic is one of the few things in geometrical optics that has any physical reality. Wavefronts and rays are not realizable; they are just convenient symbols on which we can hang our ideas. The caustic on the other hand is real and becomes visible by blowing a cloud of smoke in the region of the focus of a lens'. This section is concerned with two ray families that can be associated with mirages, with the aim of clarifying the connection (actually it will be a lack of connection) between their caustics and the mirage reflection.

The rays to be considered are those of the form (2.12), corresponding to the model index profile (2.11). Each ray is parametrized by the coordinates x_c, z_c of its reflection point, whose specification determines the family being considered. Thus we write, explicitly,

$$z(x;\ z_c, x_c) = z_c + \frac{(x - x_c)^2}{4z_c}. \tag{4.1}$$

For each family, the envelope (caustic) is the curve touched by all the rays.

The first family consists of all rays reflecting at the same height z_c, so the single parameter is x_c. For this translationally invariant family, the envelope is determined by

$$\frac{\partial}{\partial z_c} z(x; z_c, x_c) = 0, \text{i.e.}\, x = x_c, \quad \text{so} \quad z_{\text{caustic}} = z_c. \tag{4.2}$$

This family of rays and its caustic are shown in figure 5(a). It would be generated by a beam of parallel rays in the region of uniform index, incident on the air in which the index is increasing upwards. It is obvious that in this case the caustic coincides with the locus of reflection points. And it closely resembles the case studied in detail experimentally by Raman and Pancharatnam [1], in which the incident light consists of the parallel rays in a sunbeam.

However, this association between caustics and the reflection locus is a special feature of the particular family (4.2). In general, they are different geometrical objects. The second family illustrates this. It consists of rays emitted from a source at $x = 0$, $z = h$. Thus x_c and z_c are related by

$$h = z_c + \frac{x_c^2}{4z_c} \Rightarrow x_c(z_c) = \pm 2\sqrt{z_c(h - z_c)}. \tag{4.3}$$

Now the envelope is determined by

$$\frac{\mathrm{d}z(x; z_c, x_c(z_c))}{\mathrm{d}z_c} = \frac{\mathrm{d}}{\mathrm{d}z_c}\left(z_c + \frac{(x - x_c(z_c))^2}{4z_c}\right) = 0. \tag{4.4}$$

A little algebra gives

$$z_c = \frac{hx^2}{4h^2 + x^2}, \tag{4.5}$$

and so, after substituting into (4.2) the equation of the caustic is

$$z_{\text{caustic}}(x) = \frac{x^2}{4h}. \tag{4.6}$$

Thus for this family the envelope of all the parabolic rays (4.1) is itself a parabola, illustrated in figure 5(b). Introducing the third dimension, it is a paraboloid of revolution. This is the upside-down analogue of the gravitational ‘bounding paraboloid’, representing the limit of the region that can be reached by projectiles fired in different directions from the same source and with the same speed.

The reflection locus consists of the lowest points $z = z_c$, $x = x_c$ on all the rays. Substituting into (4.3) gives

$$\text{reflection locus} : 4\left(z - \tfrac{1}{2}h\right)^2 + x^2 = h^2. \tag{4.7}$$

This locus, also shown in figure 5(b), is an ellipse—or, in three dimensions, an ellipsoid of revolution, coinciding with the paraboloidal caustic only at its lowest point.

Families of rays, and their associated caustics, represent light emitted from a point source (cf figure 5(b)). A distant mirage-viewed scene consists of an extended source, that is infinitely many points, with a family of rays emanating from each. For some points of the scene, the caustics lie below the eye; there are two images of such points, one upright and one inverted, corresponding to the two rays passing each point above the caustic in figure 5(b). Other points, whose caustics lie above the eye, are not mirage-reflected. The boundary between the two classes of points is the ‘vanishing line’, consisting of points in the scene whose caustic passes through the eye; this line separates heights that are doubly mirage-reflected from heights that are invisible in the mirage.

5. Wave mirages

Monochromatic light waves with wave number k, in the index profile (2.11) are (with polarization neglected), solutions of the scalar Helmholtz equation

$$\nabla^2\Psi + k^2 Az\,\Psi = 0. \tag{5.1}$$

Different solutions correspond to different families of geometrical rays.

Raman and Pancharatnam [1] consider the translationally invariant wave for which the solution is separable in Cartesian coordinates, namely

$$\Psi(x, z) = \exp(\mathrm{i}k\mu_c x)\psi(z), \tag{5.2}$$

with $\psi(z)$ satisfying

$$\partial_z^2\psi + k^2 A(z - z_c)\psi = 0. \tag{5.3}$$

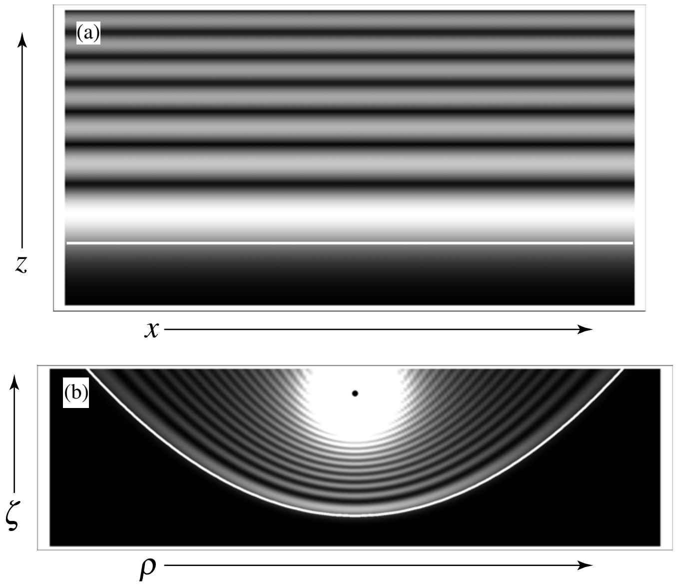

Figure 6. Intensity patterns $|\Psi|^2$ of mirage waves: (a) translationally invariant wave (5.4), corresponding the the ray family of figure 5(a); (b) wave (5.10) from a point source (the black dot), corresponding to the ray family of figure 5(b), computed for dimensionless source height $\eta = (k^2\partial_z\mu)^{1/3}h = 20$.

Nowadays we recognize the solution that decays in the region below the caustic, which for this case corresponds to the reflection locus z_c as the standard Airy function [21, 22]:

$$\Psi(x, z) = \exp(\mathrm{i}k\mu_c x)\mathrm{Ai}(-(k^2A)^{1/3}(z - z_c)). \tag{5.4}$$

Figure 6(a) illustrates this wave pattern, with the oscillations corresponding to the interference between waves corresponding to the two rays intersecting at every point above the caustic in figure 5(a).

Raman and Pancharatnam obtained exactly this result. But although the notation Ai had been in use at least since 1928 [23] and was firmly established by 1959, they used a representation in terms of Bessel functions of order 1/3, of which Jeffreys [24] remarked: '...Bessel functions of order 1/3 seem to have no application except to provide an inconvenient way of expressing this (i.e. the Airy) function'. I agree with Jeffreys, but note a contrary opinion expressed by the otherwise impeccable Dingle [25]: 'Airy functions undoubtedly possess descriptive and analytical merit.... But Bessel functions of diverse orders have an enormously wider field of application...Considered numerically, the Airy functions are non-participating and redundant anomalies liable to be superseded for numerical purposes'. History has not supported Dingle's judgement: nowadays, Ai(z) is widely available in numerical packages to arbitrary accuracy for arbitrary complex z.

Stranger still, although noting the connection with caustics Raman and Pancharatnam did not mention that the same function (though not the same notation) had been discovered by Airy in 1838 [9] precisely to describe light waves near caustics, in an analysis that was more general in that it was not restricted to this particular mirage caustic.

The second mirage wave is the solution of (5.1) corresponding to a source of waves with free-space wavenumber k at $z = h, x = y = 0$. In three dimensions, this wave depends on z and the cylindrical radial coordinate $R = \sqrt{x^2 + y^2}$, and distance from the source is

$$D = \sqrt{R^2 + (z-h)^2}. \tag{5.5}$$

The corresponding wave is the same as the time-independent Green function for a quantum particle, in a uniform electric field directed upwards. To find it, it is convenient to introduce scaled space variables, defined by

$$(R, z, h) \equiv \frac{1}{(k^2 A)^{1/3}}(\rho, \zeta, \eta). \tag{5.6}$$

In terms of these variables, the caustic (4.6) is

$$\zeta_{\text{caustic}} = \frac{\rho^2}{4\eta}. \tag{5.7}$$

The wave $\psi(\rho, \zeta; \eta)$, satisfying the Helmholtz equation (5.1) in the scaled variables with a unit source, and incorporating rotational symmetry about the z axis, is determined by the Hamiltonian equation

$$\begin{aligned} H\psi(\rho, \zeta; \eta) &= \left[-\frac{1}{\rho}\partial_\rho \rho \partial_\rho - \partial_\zeta^2 - \zeta\right]\psi(\rho, \zeta; \eta) \\ &= \frac{1}{2\pi\rho}\delta(\rho)\delta(\zeta - \eta) \end{aligned} \tag{5.8}$$

(with the definition $\int_0^\infty \mathrm{d}\rho\,\delta(\rho) = 1$). This Green function can be found in closed form. The argument, given in the appendix, is an adaptation of the corresponding Green function for one dimension, apparently first derived by Moyer [26] (see also p71 of [22]). In terms of the two parameters

$$a \equiv \zeta + \eta, \quad b \equiv \sqrt{\rho^2 + (\zeta - \eta)^2}, \tag{5.9}$$

the result is

$$\psi(\rho, \zeta; \eta) = -\frac{\mathrm{i}}{2b}\frac{\partial}{\partial b}\left[\mathrm{Ai}\left(-\frac{1}{2}(a-b)\right)\left(\mathrm{Ai}\left(-\frac{1}{2}(a+b)\right) - \mathrm{i}\,\mathrm{Bi}\left(-\frac{1}{2}(a+b)\right)\right)\right]. \tag{5.10}$$

Here Bi denotes the complementary Airy function [21]. This solution is depicted in figure 6(b).

The arguments in the Ai and Bi functions have a physical interpretation. As is clear from figure 5, each point ρ, ζ inside the caustic (i.e. for $\zeta < \zeta_{\text{caustic}}$ in (5.7)) is reached by two rays from the source. The rays are characterized by their optical path lengths, equal to the actions A_+ and A_- along them. These actions are the integrals, along the rays, of the wave vectors κ (momenta) associated with the Hamiltonian in (5.8), namely (bearing in mind the rotational symmetry)

$$H = \kappa_\rho^2 + \kappa_\zeta^2 - \zeta. \tag{5.11}$$

Here we are interested in the zero-energy trajectories (for our index profile (2.11), changing energy simply corresponds to changing the depth at which the index is zero), so the actions are

$$A = \int_0^\rho \mathrm{d}\rho' \kappa_\rho + \int_\eta^\zeta \mathrm{d}\zeta' \kappa_\zeta(\zeta'), \tag{5.12}$$

where the horizontal momentum κ_ρ is constant and the vertical momentum κ_ζ depends on ζ. In terms of the two angles $\theta_\pm$ with which rays reach the point ρ, ζ, the actions are

$$\begin{aligned} A_\pm &= \sqrt{\eta}\rho\cos\theta_\pm + \int_\eta^\zeta \mathrm{d}\zeta'\sqrt{\zeta' - \eta\cos^2\theta_\pm} \\ &= \sqrt{\eta}\rho\cos\theta_\pm + \frac{2}{3}((\zeta - \eta\cos^2\theta_\pm)^{3/2} - \eta^{3/2}\sin^3\theta_\pm). \end{aligned} \tag{5.13}$$

The angles are determined by the ray equations, which lead to

$$\tan\theta_\pm = -\frac{2}{\rho}\left(\eta \pm \sqrt{\eta}\sqrt{\zeta - \frac{\rho^2}{4\eta}}\right). \tag{5.14}$$

Then it follows after some algebra that the arguments in (5.10) are

$$\tfrac{1}{2}(a-b) = \left[\tfrac{3}{4}(A_+ - A_-)\right]^{2/3}, \quad \tfrac{1}{2}(a+b) = \left[\tfrac{3}{4}(A_+ + A_-)\right]^{2/3}. \tag{5.15}$$

With these identifications, asymptotics of (5.10) above the caustic, using the known trigonometric approximations for the Airy functions (equations (9.7.9) and (9.7.11) of [21]) gives ψ as the sum of two complex exponentials, each associated with one of the two contributing rays, whose interference gives the oscillations inside the caustic, visible in figure 6(b). The corresponding phases are the individual actions A_+ and A_-, exactly as expected from the simplest wave-decorated geometrical-optics approximation. Without this approximation, the first Airy function in (5.10), with argument $-\frac{1}{2}(a-b)$, describes the universal behaviour across any smooth caustic. Usually this is the result of the procedure of uniform approximation (described for example in section 36.12 of [21]), but for the present problem it is exact.

As already stated, the function (5.10), upside-down, also represents the energy-dependent Green function for a quantum particle in a linear potential, that is, a uniform force. I considered this earlier [27], as a representation of the wave associated with 'gravity's rainbow' for a spray of monochromatic neutrons falling after being emitted from a localized source, but was unaware of the closed form (5.10). This result [22, 26], apparently not widely known even in one dimension, is a precious addition to the rather small list of exactly-solvable quantum problems.

6. Concluding remarks

There are several ways in which the foregoing could be useful in the teaching of theoretical optics at advanced undergraduate or graduate level. First, for the derivation (in section 2) of rays bending in a continuously varying stratified medium based on the conserved quantity that follows from Snel's law of refraction. The argument is standard [10, 11], but not usually presented to optics students.

Second, for the limiting argument (in section 3). The interest of this is not only as a slightly unusual example of light transmission in a multilayered medium [17, 28], but more generally, as an example of a limiting process with an unexpected singularity. Such singular limits are common in physics [29, 30] but are not usually emphasized.

Third, as a gentle way of introducing ray caustics, as fundamental and prominent features of families of rays. Even now, when the mathematics and physics of caustics are well understood [31, 32], they are surrounded by an aroma of exoticism. One reason could be that the explicit calculation of the shapes of caustics often requires tricky algebra, involving awkward implicit functions. For the two examples in section 4, the calculations are very simple.

Fourth, as a simple example of the Airy function describing waves decorating smooth caustics. The Airy function is universal, and Airy's original argument [9], involving diffraction integrals, is not very complicated, but demonstrating its generality requires forbidding mathematics [33]. By contrast, the first of the two cases discussed in section 5 involves a straightforward solution of a differential equation in terms of a standard special function [21, 34]. Usually this solution appears as in the context of the quantum physics of particles in a linear potential, but students will surely appreciate seeing the Airy function as an aspect of the mirage phenomenon that many of them will have seen. Moreover, it gives a different perspective on the central role of the Airy function in the rainbow [9, 35], and the recent revival of interest in curved caustics ('Airy beams') [36], following the observation [37] that in time-dependent quantum mechanics (analogous to paraxial optics) the Airy function describes a freely moving 'wave packet' which accelerates and does not spread.

Fifth, as an example (the second in section 5, and the appendix) of the explicit calculation and understanding of a time-independent Green function.

Sixth, to illustrate the interesting mistakes that first-rate scientists can make. More than four decades after the deaths of Raman (in 1970) and Pancharatnam (prematurely, in 1969), it is hard to assess the balance of contributions to their study [1] of the mirage—their only joint paper, it seems. But Pancharatnam's collected works [38, 39] and a detailed scientific biography of Raman [40, 41], lead me to conjecture that the confusion about geometrical optics was Raman's, the correct wave analysis, and rediscovery of the Airy function, was Pancharatnam's, and the ingenious experiments they report were conducted jointly.

Acknowledgments

I am grateful to Professor Eugeny Abramochkin for showing me the result (A.6) and helping me to understand it, Professor Andrew T Young for careful comments and very helpful historical insights and suggestions, Dr Mark Dennis for several suggestions, and a referee for suggesting the discussion of the vanishing line at the end of section 4. My research is supported by the Leverhulme Trust.

Appendix. Derivation of mirage wave (5.10) from source

Exploiting the quantum analogy, the solution of equation (5.8) can be expressed as an integral over the corresponding time-dependent Green function $\Phi(\rho, \zeta, \tau; \eta)$. As can be confirmed by substitution,

$$\psi(\rho, \zeta; \eta) = \mathrm{i} \int_0^\infty \mathrm{d}\tau\, \Phi(\rho, \zeta, \tau; \eta), \tag{A.1}$$

where Φ satisfies the time-dependent Schrödinger equation for a wave concentrated at the source at time $\tau = 0$:

$$\begin{aligned} &H\Phi(\rho, \zeta, \tau; \eta) = \mathrm{i}\partial_\tau \Phi(\rho, \zeta, \tau; \eta), \\ &\Phi(\rho, \zeta, 0; \eta) = \frac{1}{2\pi\rho}\delta(\rho)\delta(\zeta - \eta), \end{aligned} \tag{A.2}$$

where H is the operator in (5.8). Symbolically, with the waves ψ and Φ represented by the operators $\boldsymbol{\psi}$ and $\boldsymbol{\Phi}(\tau)$, this transformation corresponds to

$$H\boldsymbol{\psi} = 1 \Rightarrow \boldsymbol{\psi} = \frac{1}{H} = \mathrm{i}\int_0^\infty \mathrm{d}\tau \exp(-\mathrm{i}H\tau) = \mathrm{i}\int_0^\infty \mathrm{d}\tau\, \boldsymbol{\Phi}(\tau). \tag{A.3}$$

The exact solution of (A.2), which again can be confirmed by direct substitution, is

$$\Phi(\rho, \zeta, \tau; \eta) = \frac{1}{(4\pi\mathrm{i}\tau)^{3/2}} \exp\left\{\mathrm{i}\left(-\frac{\tau^3}{12} + \frac{\rho^2 + (\zeta - \eta)^2}{4\tau} + \frac{\tau(\zeta + \eta)}{2}\right)\right\}. \tag{A.4}$$

The phase in the exponent is the time-dependent action for particles emitted from $\rho = 0$, $\zeta = \eta$ at $\tau = 0$ and reaching ρ, ζ at time τ. For $t \ll 1$, only the middle term in the exponent is significant, and integration over space confirms the unit δ source in (A.2).

Substitution into (A.1) and using the definitions (5.9) gives the desired wave as

$$\psi(\rho,\zeta;\eta) = \frac{1}{4(\mathrm{i}\pi)^{3/2}}\frac{1}{b}\frac{\partial}{\partial b}\int_0^\infty \frac{\mathrm{d}\tau}{\tau^{1/2}}\exp\left\{\mathrm{i}\left(-\frac{\tau^3}{12}+\frac{b^2}{4\tau}+\frac{a\tau}{2}\right)\right\}. \tag{A.5}$$

The integral (in which the denominator is $\tau^{-1/2}$ rather than $\tau^{-3/2}$) corresponds to the one-dimensional Green function for a particle in a uniform electric field. It was evaluated by Moyer [26] (see also p71 of [22]), who showed that

$$\begin{aligned}\int_0^\infty \frac{\mathrm{d}\tau}{\tau^{1/2}}\exp\left\{\mathrm{i}\left(-\frac{\tau^3}{12}+\frac{b^2}{4\tau}+\frac{a\tau}{2}\right)\right\} &= 2\pi^{3/2}\exp\left(\frac{1}{4}\mathrm{i}\pi\right)\\ &\times\left[\mathrm{Ai}\left(-\frac{1}{2}(a-b)\right)\left(\mathrm{Ai}\left(-\frac{1}{2}(a+b)\right)-\mathrm{i}\,\mathrm{Bi}\left(-\frac{1}{2}(a+b)\right)\right)\right]\end{aligned} \tag{A.6}$$

from which (5.10) follows immediately.

That the solution (5.10) satisfies the source condition in (5.8) can be confirmed by using the fact that close to the source the variable b is small (cf (5.9)), so Ai and Bi can be expanded to first order in b. After taking the derivative in (5.10), the dependence on a disappears after using the Wronskian relation between Ai and Bi (see. e.g. (9.2.7) of [21]), leaving

$$\psi(\rho,\zeta;\eta)\to\frac{1}{4\pi b}=\frac{1}{4\pi\sqrt{\rho^2+(\zeta-\eta)^2}}\quad(\rho\to 0,\ \zeta\to\eta). \tag{A.7}$$

Now, integrating over a small sphere S surrounding the source gives, using an obvious notation,

$$\begin{aligned}\iiint_S \mathrm{d}^3\rho H\psi &= -\iiint_S \mathrm{d}^3\rho(\nabla^2+\zeta)\psi\\ &= -\frac{1}{4\pi}\iint_{\partial S}\mathrm{d}^2\rho n\cdot\nabla\frac{1}{b}=\frac{1}{4\pi}\iint_{\partial S}\mathrm{d}^2\rho\frac{1}{b^2}=1.\end{aligned} \tag{A.8}$$

References

[1] Raman C V and Pancharatnam S 1959 The optics of mirages *Proc. Ind. Acad. Sci.* **A49** 251–61
[2] Raman C V 1959 The optics of mirages *Curr. Sci.* **29** 309–13
[3] Thomson J 1873 On atmospheric refraction of inclined rays, and on the path of a level ray *British Association for the Advancement of Science Report* 1872 vol 42 pp 41–5
[4] Bravais A 1856 Explication, par le système des ondes, d'un cas remarquable de la réfraction de la lumière *Ann. Chim. Phys.* **46** 492–501
[5] Everett J D 1873 On the optics of mirage *Phil. Mag.* **45** 161–72
[6] Vince R S 1799 The Bakerian lecture. Observations upon an unusual horizontal refraction of the air; with remarks on the variations to which the lower parts of the atmosphere are sometimes subject *Phil. Trans. R. Soc. Lond.* **89** 13–23
[7] Wollaston W H 1800 On double images caused by atmospheric refraction *Phil. Trans. R. Soc. Lond.* **90** 239–54
[8] Biot J B 1810 *Recherches Sur Les Réfractions Extraordinaires Qui Ont Lieu Près De L'Horizon* (Paris: Garnery)
[9] Airy G B 1838 On the intensity of light in the neighbourhood of a caustic *Trans. Camb. Phil. Soc.* **6** 379–403
[10] Adam J A 2003 Mathematics in nature: modelling patterns in the natural world (Princeton, NJ: Princeton University Press)
[11] Adam J A 2009 *A Mathematical Nature Walk* (Princeton, NJ: Princeton University Press)
[12] Minnaert M G J 1993 *Light and Color in the Outdoors* (Berlin: Springer)
[13] Greenler R 1980 *Rainbows, Halos and Glories* (Cambridge: Cambridge University Press)
[14] Lynch D K and Livingston W 1995 *Color and Light in Nature* (Cambridge: Cambridge University Press)
[15] Young A T 2012 An introduction to mirages http://mintaka.sdsu.edu/GF/mirages/mirintro.html
[16] Young A T 2012 Annotated bibliography of atmospheric refraction, mirages, green flashes, atmospheric refraction, etc http://mintaka.sdsu.edu/GF/bibliog/bibliog.html
[17] Born M and Wolf E 2005 *Principles of Optics* (London: Pergamon)
[18] Laplace M 1805 *Traité de Mécanique Céleste, Tome Quatrième* (Paris: Libraire pour les Mathématiques)
[19] Shastry G B 1978 Teaching mirages *Am. J. Phys.* **46** 765

[20] Stavroudis O N 1972 *The Optics of Rays, Wavefronts and Caustics* (New York: Academic)
[21] DLMF 2010 *NIST Handbook of Mathematical Functions* (Cambridge: Cambridge University Press) http://dlmf.nist.gov
[22] Vallée O and Soares M 2010 *Airy Functions and Applications to Physics* 2nd edn (London: Imperial College Press)
[23] Jeffreys H 1928 The effect on love waves of heterogeneity of the lower layer *Geophys. Suppl. MNRAS* **2** 101–11
[24] Jeffreys H 1942 Asymptotic solutions of linear differential equations *Phil. Mag.* **33** 451–6
[25] Dingle R B 1973 *Asymptotic Expansions: Their Derivation and Interpretation* (New York: Academic)
[26] Moyer C A 1973 On the Green function for a particle in a uniform electric field *J. Phys. A: Math. Nucl. Gen.* **6** 1461–6
[27] Berry M V 1982 Wavelength-independent fringe spacing in rainbows from falling neutrons *J. Phys. A: Math. Gen.* **15** L385–8
[28] Carron I and Ignatovich V 2003 Algorithm for preparation of multilayer systems with high critical angle of total reflection *Phys. Rev.* A **67** 043610
[29] Berry M V 1994 Asymptotics, singularities and the reduction of theories *Proc. 9th Int. Cong. Logic, Method., and Phil. of Sci.* ed D Prawitz, B Skyrms and D Westerståhl (Amsterdam: Elsevier) pp 597–607
[30] Berry M V 2002 Singular limits *Phys. Today* **55** 10–11
[31] Berry M V and Upstill C 1980 Catastrophe optics: morphologies of caustics and their diffraction patterns *Prog. Opt.* **18** 257–346
[32] Nye J F 1999 *Natural Focusing and Fine Structure of Light: Caustics and Wave Dislocations* (Bristol: Institute of Physics Publishing)
[33] Ludwig D 1966 Uniform asymptotic expansions at a caustic *Commun. Pure Appl. Math.* **19** 215–50
[34] Berry M V 2001 Why are special functions special? *Phys. Today* **54** 11–2
[35] Lee R and Fraser A 2001 *The Rainbow Bridge: Rainbows in Art, Myth and Science* (Bellingham, WA: Pennsylvania State University and SPIE Press)
[36] Siviloglou G A, Broky J, Dogariu A and Christodoulides D N 2007 Observation of accelerating Airy beams *Phys. Rev. Lett.* **99** 213901
[37] Berry M V and Balazs N L 1979 Nonspreading wave packets *Am. J. Phys.* **4** 264–7
[38] Series G W 1975 *Collected Works of S. Pancharatnam* (Oxford: Oxford University Press)
[39] Berry M V 1994 Pancharatnam, virtuoso of the Poincaré sphere: an appreciation *Curr. Sci.* **67** 220–3
[40] Venkataraman G 1988 *Journey Into Light: Life and Science of CV Raman* (Bangalore: Indian Academy of Sciences)
[41] Berry M V 1989 Review of 'Journey into light: life and science of C V Raman by G Venkataraman' *Nature* **338** 685–6