Mirage in the laboratory

E. Fabri
Istituto di Astronomia, Università di Pisa, Pisa, Italy

G. Fiorio
Istituto di Elaborazione dell'Informazione del C.N.R., Pisa, Italy

F. Lazzeri and P. Violino
Istituto di Fisica, Università di Pisa, Pisa, Italy

(Received 16 February 1981; accepted for publication 13 August 1981)

The theory of the mirage is reviewed at several levels of complexity, in order to show that the phenomenon can usefully be discussed in teaching. In addition, an experimental device has been set up to obtain the mirage in the laboratory under controlled conditions, with the possibility of observing the finer details discussed in the theory.

INTRODUCTION

The mirage is a popular phenomenon whose real meaning is often misunderstood. It is often mistaken for something that does not exist, rather more a hallucination than a physical phenomenon. Nevertheless it could very well be used in teaching optics, when discussing the light refraction, since it is something different from the most usual cases of refraction (lenses). It is, however, seldom discussed in high schools because it can be observed only in rather exceptional weather conditions, and is therefore unsuitable for a systematic observation, which is highly desirable for a useful classroom presentation.

In this work we describe a simple experimental device whose purpose is to allow the observation of a mirage in controlled conditions and irrespective of the weather. In addition, it is possible to observe finer details of the phenomenon, to obtain a photographic record, and to observe different types of patterns.

As is well known, the mirage can be observed near surfaces having a different temperature from the surrounding atmosphere, thus in the presence of a temperature gradient and hence of a refractive-index gradient. By far the most common case of mirage, the one all students have probably noticed, is above a hard-surface road under the sun: the temperature gradient there is downwards since the dark pavement is hotter than the surrounding air and the light rays are curved upwards. One thus observes an additional image that appears to be below the true position of the object; such a phenomenon (that can also be observed over deserts) is therefore usually called the "inferior mirage" to distinguish it from the case when the thermal gradient is upwards. In the latter case it is sometimes possible to observe the "superior mirage" (so called because the image associated with the mirage appears above the object's true position): the light rays are curved downwards; it can be observed (much more rarely than the former case) offshore when the water temperature is much lower than the air temperature. The two phenomena, although physically very similar, give unlike sensations: in one case the curvature of the light rays is perceived as an area of still water on the ground, in the other case (since it is virtually impossible to recognize the reflected sea surface against a luminous sky), the phenomenon appears by noticing a ship that is traveling in the sky, possibly upside down. This difference springs out of the way in which our brain perceives things, and can be a useful starting point to discuss with the students the relationship between the physical world and our way to perceive it.

In this article, after a short review of the most common approach to the mirage, we shall discuss its physical aspects using a model, and the device for its laboratory observation. The theory is equally applicable to both inferior and superior mirages, even though all examples will refer to the inferior mirage only.

There are more complex cases of mirage, in the presence of an alternating thermal gradient, like the "*fata morgana*," that we shall not discuss. A nice qualitative discussion of these and many more optical phenomena in the atmosphere can be found in Ref. 1, while a more recent treatment has been given, e.g., by Greenler.[2]

ELEMENTARY THEORY

The propagation of a light ray in an isotropic inhomogeneous medium is determined by the spatial dependence of the refractive index n. If the absorption is negligible, only the real part of n has to be considered. It is $n = c\sqrt{\epsilon\mu}$. When $\mu \simeq \mu_0$ (as in the case in air), it is $n = \sqrt{\epsilon_r}$ ($\epsilon_r = \epsilon/\epsilon_0$). Within a good approximation, it is also

$$\epsilon_r - 1 = N\alpha,$$

where N is the number density of molecules and α is the molecular polarizability. Finally, as a consequence of the gas laws, in air at constant pressure, $N \propto 1/T$, and therefore it is easy to write the refractive index as a function of temperature:

$$n^2 - 1 = (n_s^2 - 1)T_s/T;$$

n and n_s being the refractive indeces at the temperatures T and T_s, respectively.

Let us consider a region of space where the temperature is constant on parallel planes, and therefore the thermal gradient is everywhere in the same direction. With reference to Fig. 1, let y be the gradient axis (positive sense for growing n). A light ray forming an angle ϑ with the y axis undergoes a series of infinitesimal refractions; $n \sin\vartheta$ being always constant. Therefore (using the subscript A for all quantities referred to the point A) all along the ray it is $n \sin\vartheta = n_A \sin\vartheta_A$. Therefore, when $n = n_A \sin\vartheta_A$, it follows that $\sin\vartheta = 1$ and the ray cannot penetrate at lower values of n; it is a concave curve, pointing on both ends towards the positive sense of the y axis. If the region with a thermal gradient is not unlimited, but bounded between

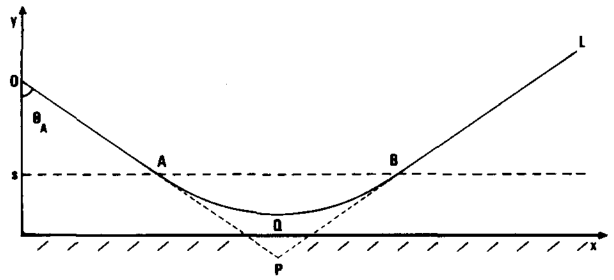

Fig. 1. Mirage: a light ray *LBQAO* from the light source *L* is bent in the layer with a thermal gradient, of thickness *s*, and the observer *O* sees it as if it had undergone a specular reflection at the point *P*. The vertical scale is much exaggerated (here and in the next figures) for sake of clarity.

two y values, $y_1 = s$ and $y_2 = 0$ (where the value of n is, respectively, n_1 and n_2; $n_1 > n_2$) it follows that a ray entering this region from $y > s$ with an angle of incidence ϑ_1 crosses the region and reaches the $y = 0$ plane if $n_1 \sin \vartheta_1 < n_2$, while in the opposite case it reaches a condition where $\sin \vartheta = 1$, and comes up again. If, therefore,

$$\sin \vartheta_1 > n_2/n_1, \tag{1}$$

an observer sees the ray emerging as if it had been reflected. For example, if $T_1 = 300$ K and $T_2 = 330$ K, in air at standard pressure, the limiting value of ϑ_1 is 89°36′. That means that an observer, at the height of 1.6 m, "sees," in these conditions, the (possibly deformed) pavement if $\vartheta <$ 89°36′, while he sees rays coming from higher regions (sky, trees, ...) if 89°36′ $< \vartheta <$ 90° (thus in grazing observation), namely, at an apparent distance $>$ 230 m.

This model is very simple and has the advantage of using only the temperature values on the boundary, and not within the region we are considering. Nevertheless, we cannot describe the shape of the ray, nor reconstruct the image of a source; it is therefore useful to also consider some more detailed model.

SPECIFIC MODEL

The temperature values near the surface are greatly affected by edge effects, namely, by the width of the surface and the characteristics of neighboring surfaces. The problem does not therefore have a general solution. In order to discuss some features of the mirage, we shall use a very simple model, which is fairly acceptable the center of the region with a thermal gradient, namely, the area mostly affecting this phenomenon.

Let us consider a layer of air of thickness s, between $y = s$ and the pavement plane $y = 0$; the temperature within this layer is a linear function of y, while it is constant for $y > s$. We have then that, for $0 < y < s$, $n^2 - 1$ is a linear function of y; therefore,

$$dn^2/dy = k, \tag{2}$$

with k a constant. If x and y are the coordinates of any point along the ray, we have

$$dy/dx = \mp \cot \vartheta$$

(the $-$ sign refers to the part of the ray where y decreases with growing x, and the $+$ sign to the other branch; this has no relevance in what follows). Then

$$\sin \vartheta = [1 + (dy/dx)^2]^{-1/2}.$$

Therefore the constant value of $n \sin \vartheta$ is

$$n[1 + (dy/dx)^2]^{-1/2} = n_m;$$

n_m being the value of the refractive index in the turning point where $dy/dx = 0$ (let $y = y_m$ there). Then

$$dy/dx = \mp (n^2 - n_m^2)^{1/2}/n_m.$$

Since n is constant above s, let n_0 be its value:

$$n = n_0 \quad \text{for } y \geqslant s;$$

integrating Eq. (2) in the interval $0 < y < s$ one obtains

$$n^2 = n_0^2 + k(y - s) \quad \text{for } 0 < y \leqslant s;$$

thus

$$dy/dx = \mp [k(y - y_m)]^{1/2}/n_m. \tag{3}$$

It is worth noting that we would have reached the same result, within a very good approximation, assuming a constant gradient of n instead of T, since the quadratic terms arising in this case would be negligible. This shows the stability of this model. Garbasso,[3] on the other hand, has shown that this model is a consequence—within the same approximation—of the thermal diffusion equation. More complex models have been considered by Hillers[4] in connection with the double reflection problem that we shall mention later.

Integrating Eq. (3) one obtains the ray equation

$$y = a + bx + cx^2, \tag{4}$$

where $c = k/4n_m^2$, while the constant b is determined by ϑ_1 and the value of a is in connection to the abscissa of the point where the ray crosses the $y = s$ plane. The ray is thus a parabola. This is true, of course, only within the region with a thermal gradient; outside it, namely, for $y > s$, the rays are straight lines.

FINER DETAILS

It is useful to express the constants a and b in Eq. (4) in terms of parameters that are measurable, at least in principle. To do this, let us assume that an observer O is along the y axis, outside the layer with thermal gradient, namely, at $y_0 > s$; let h be his distance from the $y = s$ plane (in other words, $y_0 = s + h$), and let γ be the slope ($|dy/dx|$) of a ray reaching O after being deviated in the layer with thermal gradient. It is then easy to show that the ray equation (4) becomes

$$y = s + h - \gamma x + c(h/\gamma - x)^2.$$

In this equation the parameter c is still a function of n_m; but n_m is very nearly unity, and therefore we have with negligible error

$$y = s + h - \gamma x + k(h/\gamma - x)^2/4.$$

The coordinates of the points A, B, and Q (see Fig. 1) are easily found:

$$A \equiv (h/\gamma, s),$$
$$B \equiv (h/\gamma + 4\gamma/k, s),$$
$$Q \equiv (h/\gamma + 2\gamma/k,\ s - \gamma^2/k).$$

The last equation shows that only rays with $\gamma \leqslant \surd(ks)$ can reach the observer coming from B, in agreement, always within the approximation $n \simeq 1$, with Eq. (1). Rays reaching O with $\gamma > \surd(ks)$ necessarily cross the x axis (i.e., they emerge from the pavement, and not from the $y > s$ area). The equation of the straight ray passing through B and coming from $y > s$ is

$$y = s - h - 4\gamma^2/k + \gamma x. \tag{5}$$

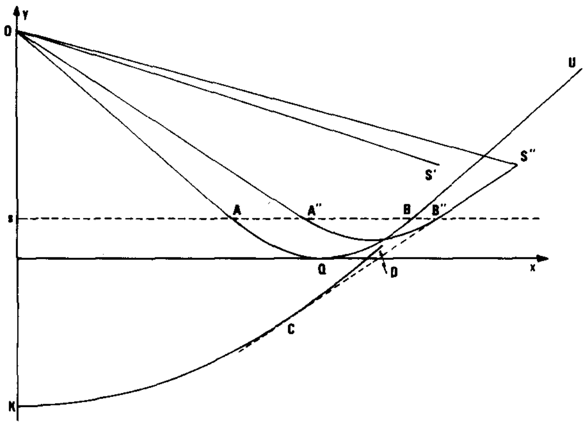

Fig. 2. Caustic KD associated with the rays reaching O and undergoing a deviation, and limiting ray $UBQAO$. For more details, see the text.

We thus have a family of straight lines when γ varies between 0 and $\sqrt{(ks)}$. It is useful considering their envelope (caustic) shown in Fig. 2. Differentiating Eq. (5) with respect to γ, one finds the abscissa of the point C where each straight line is tangent to their envelope, $x_C = 8\gamma/k$; by substitution in Eq. (5):

$$y_C = s - h + 4\gamma^2/k.$$

Getting rid of γ one obtains the equation of the envelope:

$$y = s - h + kx^2/16,$$

namely, still a parabola. The caustic, since it is bound to values $0 < \gamma < \sqrt{(ks)}$, lies between the y axis and the point D whose coordinates are

$$D \equiv (8\sqrt{(s/k)}, 5s - h).$$

We must consider two different cases: $y_D \gtrless s$, or equivalently, $h \lessgtr 4s$.

Case A: $h > 4s$. The path $OAQBU$ of the limiting ray with $\gamma = \sqrt{(ks)}$ is shown in Fig. 2; the straight line BU is tangent to the caustic in D (namely, outside the physical ray). The half-line BU divides the half-plane $y > s$ into two regions: at the left one finds points like S': no straight line tangent to the caustic can pass through it and S' can therefore be seen by O only through the straight ray $S'O$: there are no "reflections." On the contrary, through a point S'' at the right there is a tangent to the caustic and in addition to the straight ray one has a "reflected" ray $S''B''A''O$.

Case B: $h < 4s$. The limiting ray $OAQBU$ with $\gamma = \sqrt{(ks)}$ is shown in this case in Fig. 3. The caustic KHD and the half-line BU identify now in the half-plane $y > 0$ three regions: the one at the left of HD (in Fig. 3), the one between HD and BU, and the one at the right of BU. When the light source is in the first region, like S', the observer O sees no mirage, and when it is in the third region, like S'', O observes a regular mirage, exactly as in case A. With sources in the second region, like S, however, one observes a new phenomenon. Through S pass *two* different tangents to the caustic (one touching it below the pavement, as in case A, and one touching it far above the pavement); O therefore sees two different "reflected" rays, both labeled by a small S in Fig. 3, in addition to the usual ray (direct observation). In order to examine in more detail the two reflections, let us consider two different points S and T in an extended source. The reflected rays from T to O are labeled by a small T in Fig. 3. S is above T and so does it appear in direct observation; in the first reflection (the one observed in more grazing incidence) T appears to be above S: the image is thus inverted as in a mirror, like in case A. In the second reflection (in less grazing incidence) S appears again above T: the order of the two rays is the same as in the source; the image is now erect, even though it is more compressed than in direct observation.

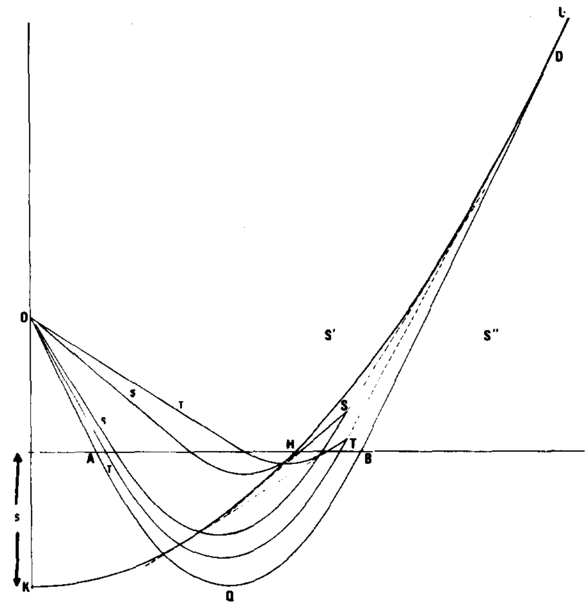

Fig. 3. Case of "double reflection." Rays from S' and S'', and straight rays SO and TO are omitted for clarity.

We can summarize by saying that upon changing the horizontal position of the source (i.e., the distance between source and observer O), at low distances (region at the left of HD in Fig. 3) there is only the direct image, then (region between HD and BU) there are both reflections (inverted and erect); at larger distances (region at the right of BU) the second reflection disappears and there is only the inverted, mirrorlike, reflection.

It is possible to find out where are the points from which

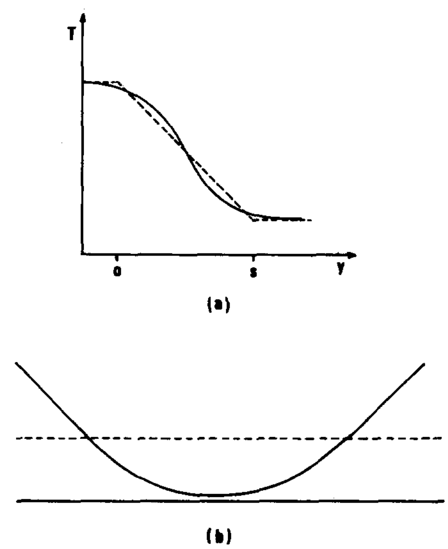

Fig. 4. When the temperature is not a linear function of y (dashed line) but has the more realistic shape shown in (a), the ray is no longer a parabola but it has rather the shape shown in (b).

the "reflected" rays appear to diverge, and whether such points coincide for sagittal and tangential bundles (in other words, whether the images are affected by astigmatism). Such calculations are not reported here since they are rather lengthy; they show, however, that the focal planes of the various images do not coincide, and that there is indeed astigmatism. This fact is of practical importance since it causes some difficulties when one attempts to photograph such images.

In a purely qualitative way we can notice that if $n^2 - 1$ is not a linear function of y throughout the $0 < y < s$ region, but has the more realistic behavior shown in Fig. 4(a), the rays penetrating more deeply are less curved (the curvature is proportional to k) and as a consequence it is easier to observe the double reflection.

EXPERIMENTAL APPARATUS

In order to study in a laboratory the above phenomena, it is necessary to build a device where the thermal gradient is obtained irrespective of weather conditions. The first such device was built by Wollaston in 1800.[5] For the observation of the double reflection, however, he used a complicated system of partially miscible fluids, whereas the first laboratory observation of this phenomenon in air (thus of greater educational interest since it it closer to the natural phenomenon) was carried out by Hillers.[6] His device consisted, on the analogy of the natural mirage, of a horizontal heated plate above which he carried out the observations. It was, however, fairly big and the image appeared fluctuating in an irregular fashion. Hillers was thus unable to obtain a photographic record of it, whereas he succeeded in "natural" conditions obtaining, among other pictures, an exceptional record of the double reflection near a 188-m-long wall along the river Elbe.[7]

In the device we developed, the heated surface is in a vertical plane, instead of a horizontal one as in Hillers's case (and therefore the y axis is horizontal). Over a heated horizontal surface ther are convective cells and turbulent motion, and the observation is thus very difficult when s is small; with a vertical surface, convection itself causes a fast, laminar air motion in front of the surface, with two simultaneous advantages: the phenomenon is stable and the thermal gradient is much higher; we could thus considerably reduce the size of the device. This reduced size, besides allowing its setup in any classroom, also allows the observations under grazing angles, keeping observer and source at a certain distance from the hot plate and making use of the optical lever.

In our device the heated vertical plate is made out of iron, its surface is smooth but not accurately machined in order to prevent reflections by the metal surface itself; its size is $86 \times 7 \times 1.2$ cm; the appreciable thickness has the purpose of preventing accidental temperature fluctuations, due, e.g., to the opening of a door. Heating is accomplished by means of five flatiron heaters attached to the plate, with

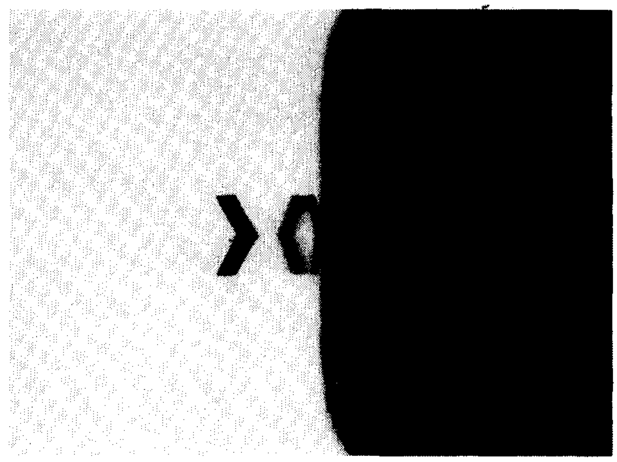

Fig. 5. Picture showing a mirage with double reflection. The heated plate is on the right; the light source was a sheet of white paper with a black sign shaped like an arrowhead. One can easily recognize (left to right) the direct image, the first reflection (inverted and slightly blurred) and the second reflection (erect and very compressed).

some thermal insulation on the opposite side in order to increase the thermal flux through the plate. The electric power delivered to the heaters can be raised up to 2 kW altogether.

If one can position the light source and the observer at a certain distance in the plane of the plate (up to 2 m each side), it is fairly easy to locate the distance y_0 from the plane of the plate best suited in order to observe the double reflection, as discussed in the previous section. A picture obtained in these circumstances is shown in Fig. 5; no reflections can be observed when the plate is at room temperature. Unfortunately, it is not easy to take good-quality pictures because the various images appear in focal planes that are rather far apart, as we mentioned before.

With this device it is easy, by changing the relative positions of the source and the observer, to study visually the conditions for the mirage to show up and its various appearances; to show the "normal" mirage and its finer details. If the electric power is delivered to the heaters by means of a variable transformer, one can also change the thermal gradient, and observe the effect of its variation.

[1]M. Minnaert, *Light and Colour in the Open Air* (Dover, London, 1940).
[2]R. Greenler, *Rainbows, Halos and Glories* (Cambridge University, Cambridge, 1980).
[3]A. Garbasso, Ann. Phys. IV **39**, 1073 (1912).
[4]W. Hillers, Phys. Z. **14**, 719 (1913); **15**, 303 (1914).
[5]This one and other historical notes can be found in Ref. 3.
[6]W. Hillers, Phys. Z. **15**, 304 (1914).
[7]W. Hillers, Phys. Z. **14**, 718 (1913).