Waves and Symmetry
Raman Centenary Symposium,
December 1988, Bangalore.

Gravitational Lenses

Rajaram Nityananda
Raman Research Institute, Bangalore 560 080, India.

The subject known as **Gravitational Lensing** is just the most recent example of a well established astronomical tradition – one studies mass in the universe by its gravitational effects, in this case on light. This survey article is meant to give people in other fields a flavour of some basic ideas, problems and results in this rather interesting area.*

In a broad sense, we are dealing with the optics of the universe, viewed as the medium through which light travels to reach astronomers. Lobachevskii, Gauss, Riemann, Clifford, and Poincaré all seem to have considered the possibility of non-Euclidean geometry describing the world we live in. This would influence the propagation of light over large distances (figure 1a). Poincaré in particular concluded that we would find it more convenient to change the laws of optics and that ". . Euclidean geometry, therefore, has nothing to fear from fresh experiments". In a sense, he has been proved right! Most discussions of gravitational lenses, including this one, use Newtonian concepts like gravitational potential and deflection appealing to the weak-field limit of general relativity for support.

Cosmological effects apart, local deviations from uniformity are the source of lensing effects. A rather subtle one was pointed out by Zeldovich and Feynman in the sixties (figure 1b). The fact that we are looking at a galaxy unobstructed implies that there is a deficiency of luminous mass in the light path compared to the overall mean. To model this simply, excavate a tube in a homogeneous universe by removing some fraction of the mass, viz. that in galaxies (figure 1b). The effect is to cause an additional divergence of rays when compared to the truly uniform universe. Density fluctuations outside the tube could then be included as a further refinement. This point surely deserves more attention than it usually receives from most textbooks and even many papers on cosmology which often present and utilise the formulae appropriate to a smoothed out universe without comment. To be fair, one does not really know exactly what fraction of the mass in the universe is in the form of discrete lumps as opposed to a smoother background.

Early papers such as that by Einstein worked out the lensing effects caused by nearby individual stars. But the subject proper was really born when Zwicky shifted attention to whole galaxies at a cosmological distance. His two points – that one would learn more about the masses of galaxies and simultaneously about the magnified distant object – continue to be the driving force today. For simplicity, both Einstein and Zwicky considered the simple but special case with an aligned source, lens, and observer and axial symmetry. In this case the so called Einstein ring image is formed (figure 1c). Image in this context simply refers to the set of all directions of rays observed at a given point. This is close to what one means by a virtual image in optics. However, there is significant astigmatism and a narrow pencil will not converge to or diverge from a single point but rather two focal lines in general. The term lens, like image, should also be understood in this broader context. Table 1 gives the relevant scales of angle, distance, mass, and time associated with lensing by galaxies. In brief, typical galaxies are expected to create multiple images

* Accordingly, references are collected in order of appearance with comments at the end. Some arguments have been put into figure captions and Table 1.

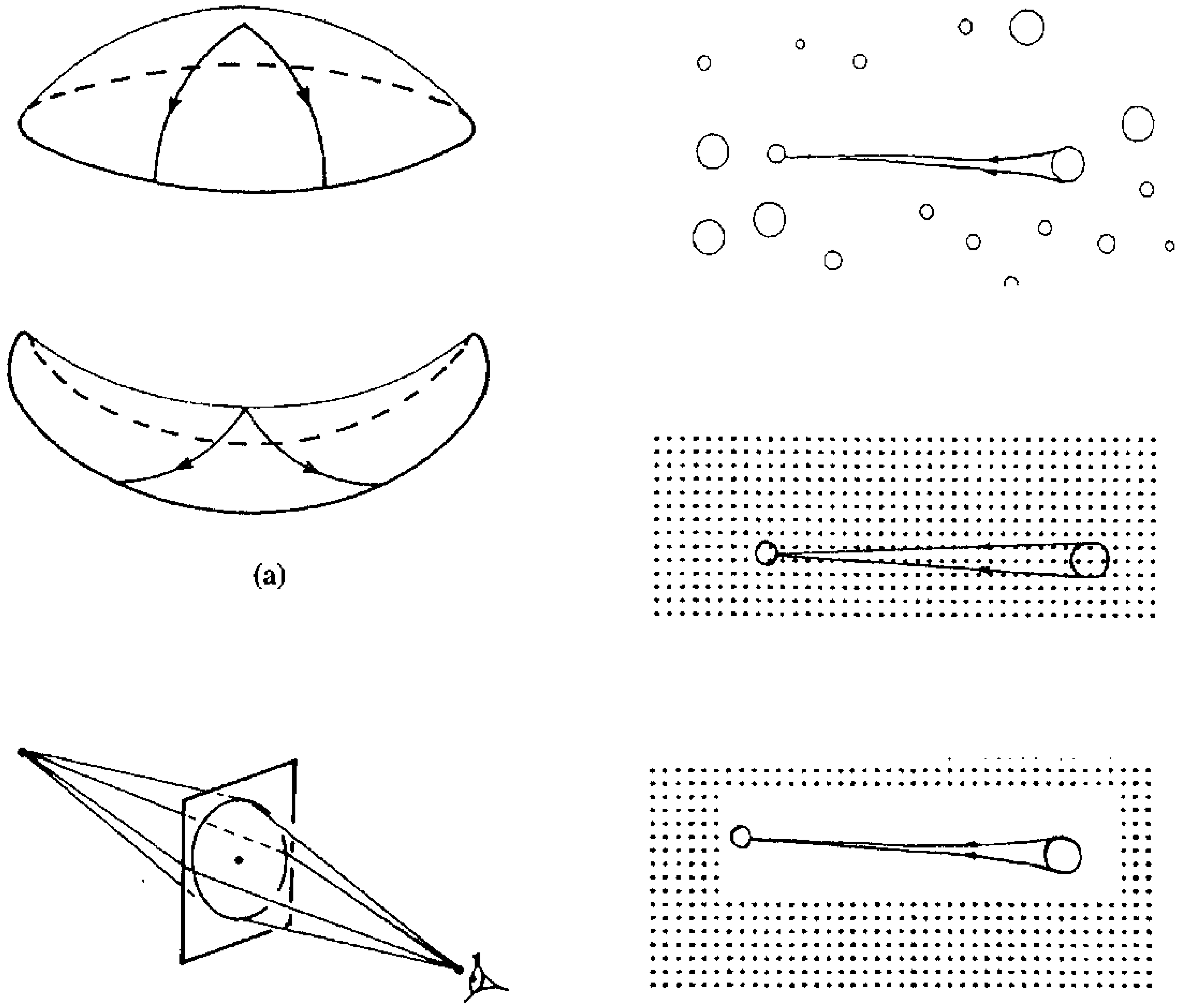

Figure 1(a). Illustrating how positive spatial curvature converges rays and negative curvature diverges them. In an expanding universe, of course, there is the additional effect that physical distances are related to those in the figure by a scale factor increasing with time. **(b)** A beam of light traversing a lumpy universe (top) suffers an extra divergence when compared to the smooth case with the same density (middle), so long as it slips between the lumps. A simple model would be a tube excavated in a uniform universe (bottom). (c) The ring image envisaged by the founding fathers of the subject occurs only in a very special symmetric geometry of aligned sources lens, and observer.

with a spacing of roughly an arc second when located at a cosmological distance. The probability of a given distant source being multiply-imaged is not too small given the large number of galaxies, and could even be close to 1 if there were dark discrete masses with a density sufficient to close the universe. Most of table 1 was known to the experts even before the discovery of the first gravitational lens in 1979 (figure 2a). This is rightly held up as a rare example of theory preceding observation in astronomy. But the observations, when they came, forced everyone to think harder.

Lensing by an axisymmetric potential is rather comfortable to handle – the symmetry forces a planar geometry and a simple graphical method gives the image positions (figure 2b). For a general smooth potential (appropriate to a galaxy) it was realised that the number of images was necessarily odd. William Burke used Poincaré's notion of the index of a vector field to show that the odd number of images was a general property (figure 2c). Two questions raised by the first lens were (i) What happened to the third image? (ii) Why is the mass inferred from the large (nearly seven arc second) image separation so much greater than for a single galaxy? Currently popular answers are (i) The potential is either nearly or actually singular giving a very weak or missing third image and (ii) dark matter, associated with a nearby cluster of galaxies, greatly enhances the deflection caused by the galaxy actually seen.

Table 1. Characteristic Scales for Gravitational Lensing by Galaxies

Quantity	Symbol or formula and order of magnitude	Remarks
1) Longitudinal distances/travel times	$D \sim 10^{28}$ cm, 10^{10} yrs	For cosmologically distant objects, the relevant timescale is the age of the universe.
2) Mass of a galaxy	$M \sim 10^{11}$ solar masses i.e. 3.10^{44} gm	Most observed cases of lensing correspond to masses larger than this.
3) Schwarzschild radius of a mass M	$R_s = 2GM / c^2$ $\quad 3 \times 10^{16}$ cm or 3×10^{-2} light years	The mass M of a galaxy is of course distributed over a much larger scale of 3×10^4 light years.
4) Impact parameter	b —	—
5) Deflection at impact parameter b (Einstein's formula for a spherically symmetric mass)	$\theta(b) = \frac{4GM}{c^2 b} = \frac{2R_s}{b}$	$M(b)$ is the mass enclosed inside a distance b and $R_s(b)$ the corresponding Schwarzschild radius.
6) Impact parameters relevant for significant magnification and multiple imaging (strong lensing)	$D\theta(b) \sim b$ i.e. $\frac{DR_s}{b} \sim b$ $b = (R_s D)^{1/2} \sim 10^4$ light years	The quantity $bD(b)$ is the amount by which the ray is deviated from its unperturbed path by the time it reaches the observer. When this is $\sim b$ the deflection is significant.
7) Surface density within b for strong lensing	$\sum = \frac{M}{R_s D} \sim \frac{c^2}{GD} \sim 1\ \mathrm{g/cm^2}$	This characteristic scale is independent of mass!
8) Angular separation of images	$\theta \sim \frac{b}{D} \sim \sqrt{\frac{R_s}{D}} < 1$ arcsec	Note that mass scales as the square of the angle. Even a few arc seconds corresponds to masses significantly greater than that of a single galaxy.
9) Cross section for lensing by a galaxy scale mass	$\sigma \sim \theta^2 \sim \frac{R_s}{D} \sim 10^{-12}$ steradian $\sim GM / c^2 D$	The solid angle surrounding the mass within which the source must be for strong lensing. Refsdal remarks that this is the gravitational potential of the lens at the observer in units of c^2.
10) Typical time delay between images	$\tau \sim \frac{D\theta^2}{c} \sim \frac{R_s}{c} \sim 2$ weeks	The quantity estimated is the geometric time delay and the gravitational one is of the same order in cases of strong lensing. Note that for $\theta \approx 6$ arc seconds this would go up to a year. Refsdal pointed out that measuring τ and θ gives D, the cosmological distance scale.

My favourite gravitational lens is the second one to be discovered (figure 5a). This lens is familiarly known either as the triple quasar or by its first name (1115). Any comfort one might have drawn from the three images originally found was shortlived since the brightest one was soon resolved as a close pair. The Pasadena and Bombay groups succeeded in finding reasonable quantitative models for the lensing geometry. Clearly the potential had to deviate from spherical symmetry since the images were not in a straight line. The best modelers surely develop a feel for what is going on in their nonlinear image finding programs. But for the rest of us, I will now give some qualitative arguments to show that this image geometry is not only understandable but perhaps even natural!

The starting point is to think not in terms of rays but wavefronts – surfaces normal to these rays. This idea really goes back to Huyghens' discussions of wave propagation. Purely as a property of rays

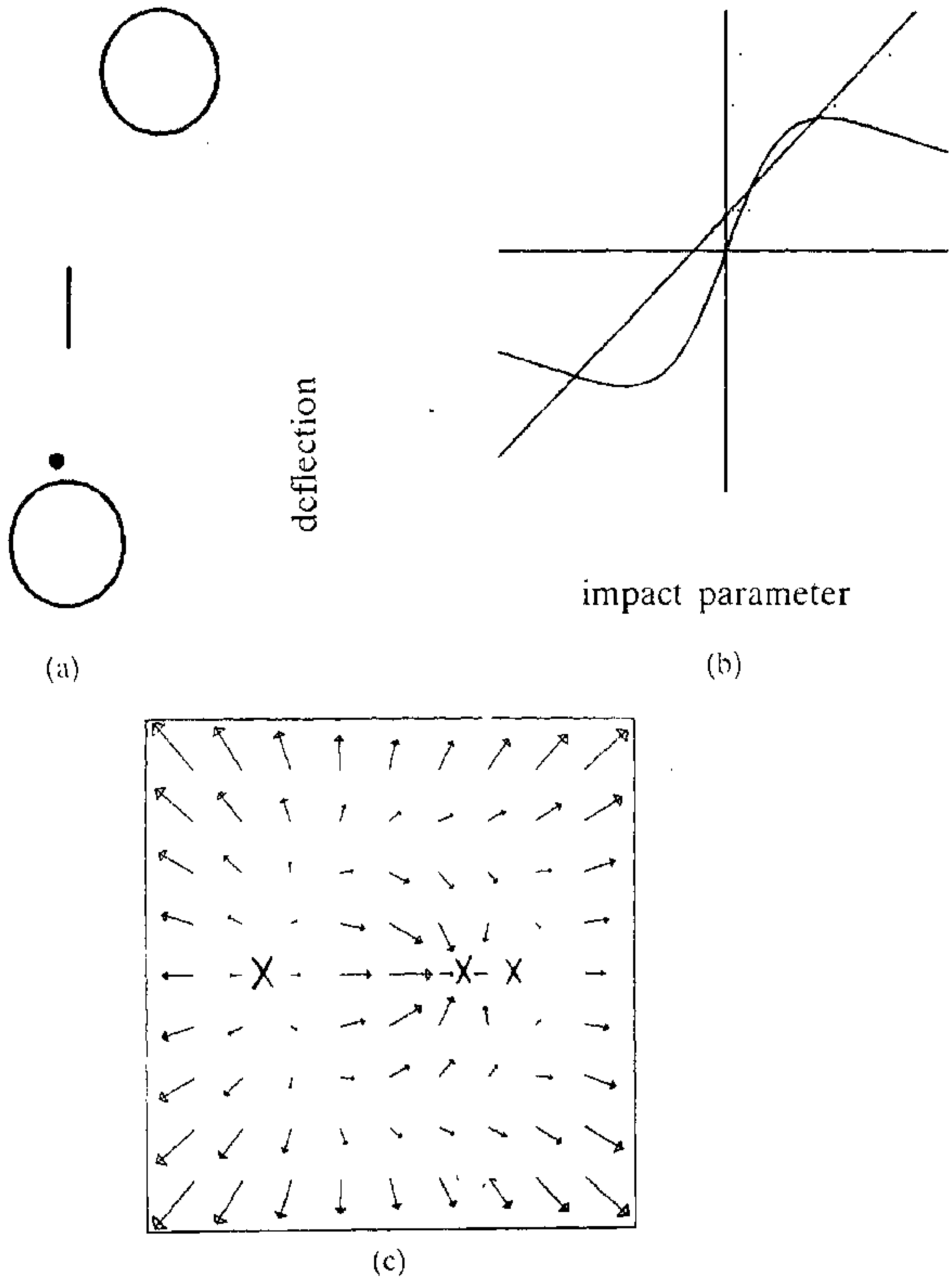

Figure 2(a). The geometry of the first gravitationally lensed pair of quasar images (often circles) to be discovered (0957+561). The superposed foreground galaxy (dot), believed to be primarily responsible, is actually fainter and was detected later. The bar is one arcsecond. **(b)** A graphical representation of the condition that the rays pass through the observer is shown. Basically, one equates two quantities (i) the deflection needed to bring a ray from impact parameter b to the observer. This is linear in b. (ii) the deflection actually produced by the lens at b (the curve). Note the odd number of intersections. **(c)** William Burke's proof of the odd number theorem is the general (two dimensional) case. At each b, the "miss vector" which defines the amount by which the ray misses the observer is plotted. At large b, this is radial and has an index + 1 i.e. rotates through 2π as we move on a contour around the origin. Images are places where the vector field vanishes. As the figure shows, these carry an index $\pm$ 1 and the total has to add up to +1, giving an odd number of images.

starting at a point and undergoing an arbitrary reflection, it was stated (much later!) by Malus but it really flowered in the care of Hamilton who made it a powerful tool not only in optics but also mechanics. If rays are to cross as in figure 3a, the corresponding pieces of wavefront must do likewise, resulting in a multi-sheeted surface which moves past the observer (figure 3b). Note that

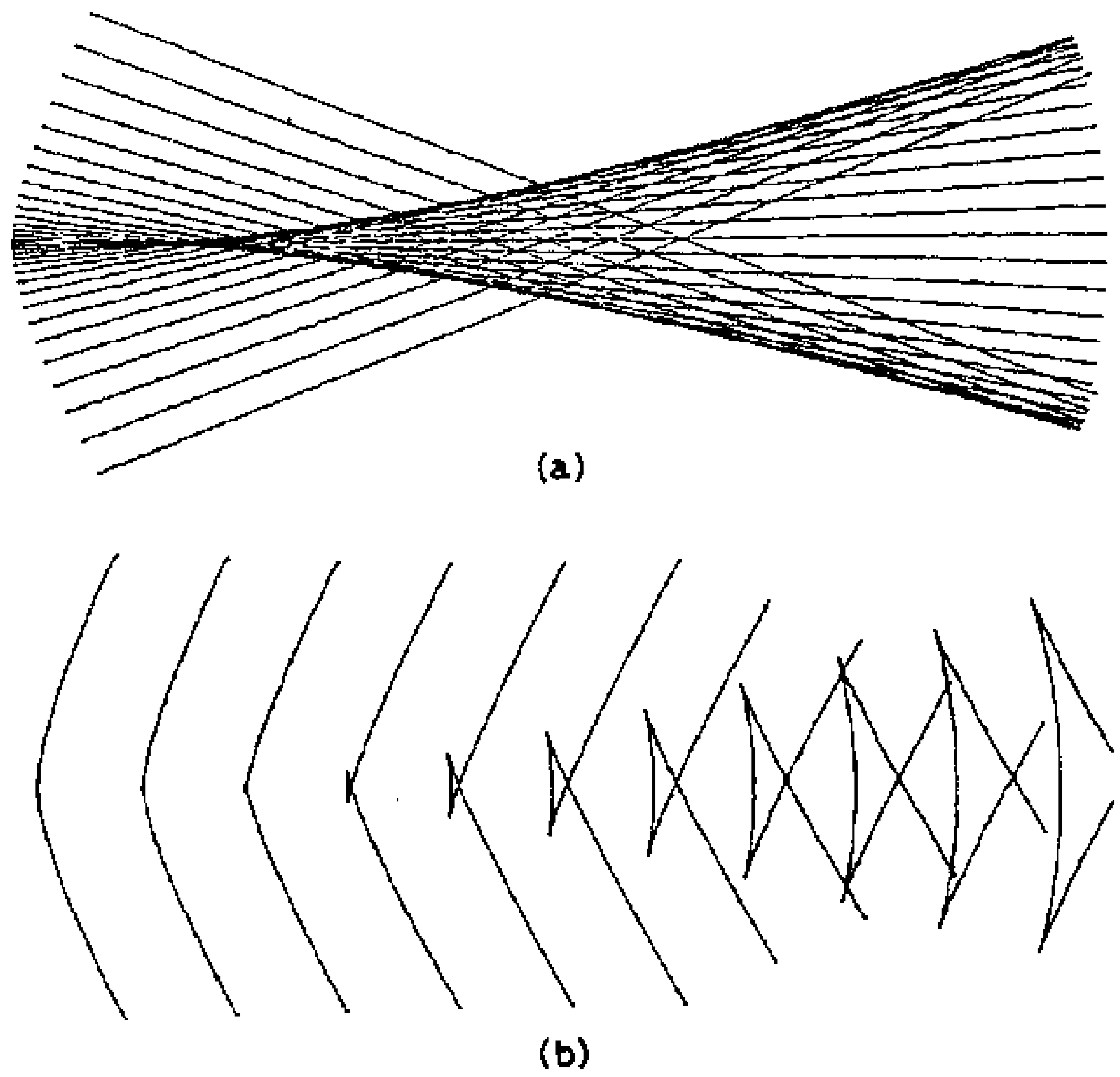

Figure 3. The relationship between rays, and caustics (a) and wavefronts (b) illustrated for a 1 + 1 dimensional case. Notice several general features (i) Multiple intersecting rays i.e. multiple images for observers within the caustic (ii) Merging of a pair of images as the observer approaches the caustic (and merging of three at the cusp) (iii) Crowding of rays corresponding to high intensities at the caustic.

there are cusps where the different sheets are attached to each other. A little thought shows that this is inevitable since all sheets have to face the direction of propagation. Different image directions correspond to the normals to the various sheets and time delays to their longitudinal separation. These remarks made by Refsdal in 1964 may look elementary but their full significance took nearly twenty years to really sink in. As figure 3 shows in the simplest case, the rays envelop caustics which are also the places where the different sheets of the wavefront are attached to each other. The great advantage of the wavefront is that it is a global concept, telling us what all observers see. This includes image magnifications which can be calculated from the principal curvatures of the wavefront for a single lens. The great disadvantage is that the function describing the wavefront is singular and multi-valued.

If one is more modest and asks for what a single observer sees, there is a smooth and single-valued function which can be used. Start with figure 4, which illustrates how gravitational deflection of rays is related to gravitational time delay of a plane wave front. The figure shows a wavefront emitted by the source which has just been deformed by the lens. The other wavefront shown is one emitted by the observer, backward in time. Points on this second wavefront will all reach the observer at the same time. The time delay can therefore be visualised as the clearance between two wavefronts. Where the normals of the two surfaces agree, we have a ray emitted by the source passing through the observer, i.e. an image. Equivalently, the two surfaces are tangent to each other at the images – the time delay has a stationary value (maximum, minimum or saddle point). Armed with the time delay function, conveniently represented by the height of a surface, one can now put in various effects. The different stages of building up an image geometry like the quadruple quasar are explained in detail in figures 5b, 5c, and 5d. One sees that four images, almost forming a ring, with two very close ones being the brightest, constitute a rather natural configuration.

Of course, there is much more to gravitational lensing than these two case studies might indicate. Unavoidably this article can only touch very briefly on many other astrophysically significant topics. One is called amplification bias. This refers to the tendency of intrinsically improbable image

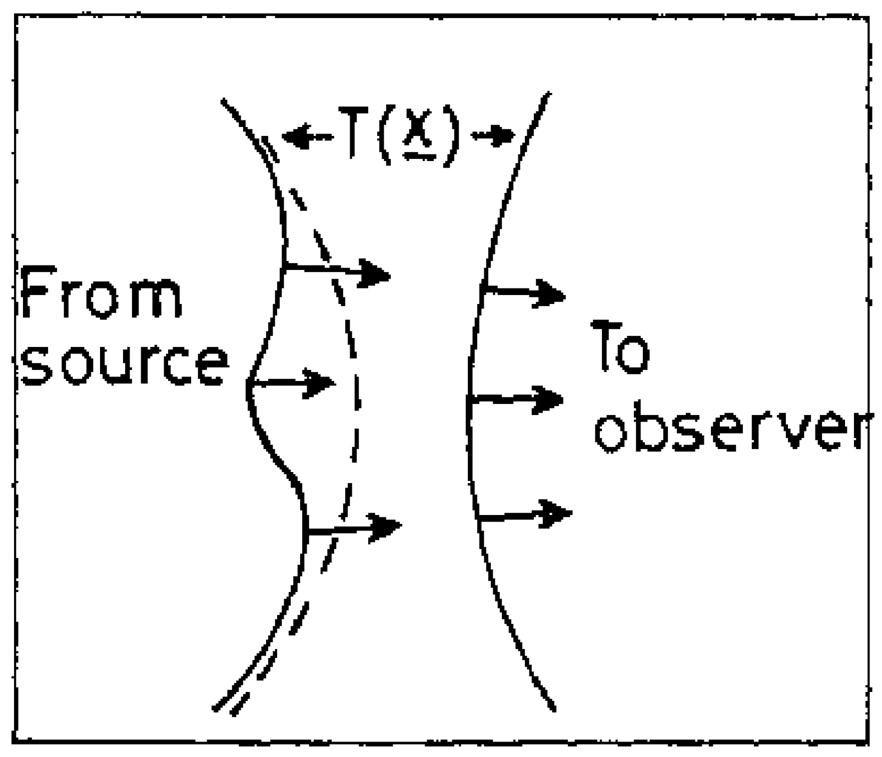

Figure 4. Geometrical construction showing a plane wavefront which has just passed the lens. Clearly one can think of the deflection as the slope of the emerging wavefront, i.e. the gradient of the gravitational time delay. This is being compared to a time reversed wavefront emitted by the observer. Those points where the normals to the two agree lie on rays which pass through the observer i.e. give the locations of images. It is convenient to plot the spacing between these two wavefronts as the arrival time function or surface. Its stationery points give the locations of images.

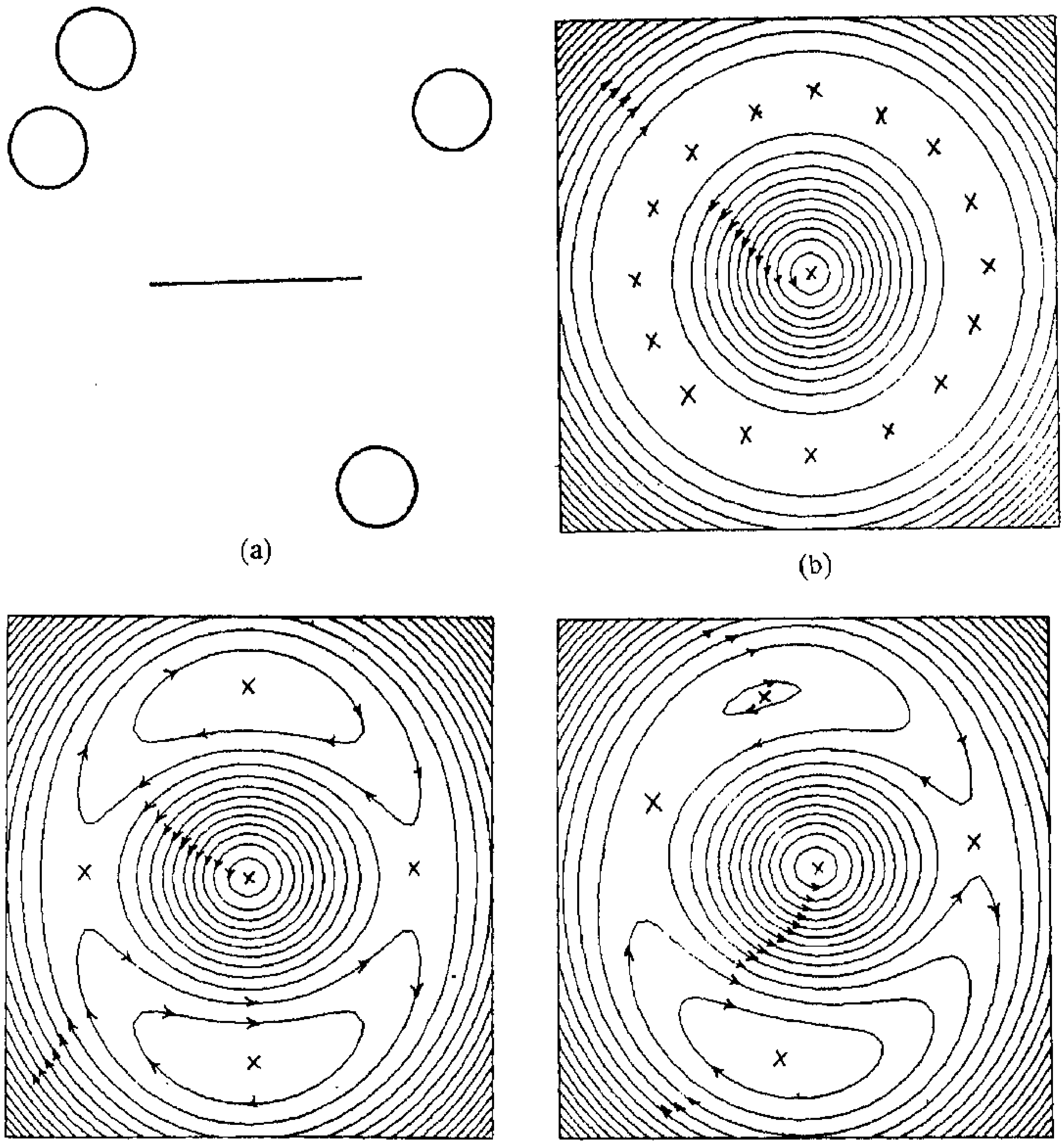

Figure 5. **(a)** Image geometry for the second gravitational lens to be discovered (1115 + 080). The close pair are much brighter than the other two. Note the one arcsecond scale bar. **(b)** Arrival time surface for the ring image geometry. The time delay has a degenerate stationary value all along the ring as shown by the crosses. **(c)** Effect of a quadrupolar distortion of the lens on the time delay along the ring. The degeneracy is broken giving four images at right angles. **(d)** Effect of a slight offset of source, lens, or observer, on the images.

configurations to be favoured in our observational sample if some images are bright (1115 is an example). The extent to which this modifies our observed picture of the distant universe is hotly debated even today.

A second topic is known as microlensing. Galaxies are after all not smooth potential wells but have stars in them. For the purpose of computing the time delay or its first derivative, the deflection, one can show that this discreteness is unimportant but for magnification which depends on the second derivative it is vital. The argument is that time delay depends on impact parameter b as $ln b$ and the weight factor is $2\pi b\, db$ giving a dominance by large values of the impact parameter. The mean deflection is again dominated by large b. However, the mean square fluctuations in the magnification are described by $(1/b^2)^2\, b\, db$ and are hence dominated by stars close to the beam. The intensities of images viewed through a dense concentration of stars such as the nucleus of a galaxy are expected to fluctuate on the timescale of many years as individual stars move closer to and away from the beam. Detecting such fluctuations in an image seen through dark matter would give direct information on the discrete objects which make up this matter. Given the velocities of the objects from dynamical modeling, the timescale of the fluctuations gives their spacing and hence, for a given density, their masses.

There is one recent observational discovery which is worth mentioning in a little more detail. Less than two years ago, mysterious arc like structures were detected in clusters of galaxies. After some amount of speculation and confusion, it became clear that these were gravitational lens images of background galaxies with the potential of the cluster acting as a lens. The scale of the arcs immediately gives important information about dark matter in clusters of galaxies – a subject also studied by Zwicky. A similar situation, from the point of view of optics, discovered at about the same time, is a radio image of a complete ring, again believed to be gravitationally lensed. This raises an interesting question, since we have argued that the ring geometry has zero probability and breaks up under the smallest perturbations. Why is it observed and what restores the symmetry? The answer is that the source is extended. Let us go back to the perfectly aligned Einstein ring geometry of figure 1c. The effect of perturbations (like moving the observer or weakly changing the potential) is that rays from the observer to the ring now miss the source. But if the source is large enough, they will just hit some other part of it and the ring image, suitably thickened, will be seen. The finite extent of the source seems to restore the symmetry.

There is some irony in presenting the subject of gravitational lenses in this meeting on waves and symmetry. Wave effects are unimportant, although wavefronts are all important. The very symmetric situations which are easiest to analyse prove to be misleading. When perturbations break the symmetry, rather beautiful patterns of wavefronts and caustics emerge which have a symmetry of their own. The universal nature of these patterns and their classification is the subject of catastrophe (or more modestly singularity) theory.

Finally, it is a pleasure to thank many people with whom I have had the privilege of discussing different aspects of this subject. In roughly chronological order they are Berry, Chitre, Narayan, Subramanian, Ostriker, Turner, Paczynski, Gott, Blandford, Padmanabhan, Kovner, and Narasimha. As the references below show, it is remarkable how many basic ideas originate from Sjur Refsdal of the Hamburg Observatory, who has rightly been described as the most durable prophet in this subject.

BIBLIOGRAPHY

Non Euclidean Geometry in physics and astronomy

The opening chapter of A. S. Eddington's book on "Space, Time, and Gravitation" has a very interesting conceptual discussion and some historical remarks.

The effect of inhomogeneities on light propagation

Zeldovich, Ya. B., *Astron. Zh.* 1964, **41,** 19 Translation: *Sov. Astron.* 1964, **8,** 13.

Gunn, J. E., *Astrophys. J.* 1967, **150,** 737 refers to Feynman's suggestion.

Early lensing papers

Einstein, E., *Science*, 1936, **84,** 506, Zwicky, F., *Phys. Rev.* 1937, **51,** 290 and 679.

Pre-discovery discussions of lensing

Refsdal, *Mon. Not. R. Astron. Soc.*, 1964, **128,** 307 on time delay between images.

Press, W. H. and Gunn, J. E., *Astrophys. J.*, 1973, **185,** 397 on lensing probabilities by a cosmologically distributed population of discrete masses.

Bourassa, R. R. and Kantowski, R., *Astrophys. J.*, 1975, **195**, 13 on nonsymmetric lenses.

Jaroszynski, M. and Paczynski, B., *Memorie della Societe Astronomica Italiana*, 1974, **45**, 673 on the notion of critical surface density.

The first lens and its aftermath

Walsh, D., Carswell, R. F. and Weymann, R. J., *Nature*, 1979, **279**, 381. The discovery paper.

Chang, K. and Refsdal, S., *Nature*, 1979, **282**, 561.

Young, P. J., *Astrophys. J.*, 1981, **244**, 756.

Gott, J. R., *Astrophys, J.*, 1981, **243**, 140.

These three papers discuss effects of discrete masses like stars on the lensing, the last pointing out the constraints on dark matter in galaxy halos.

Young, P. J., *et al., Astrophys. J.*, 1980, **241**, 507.

Dyer, C. C. and Roeder, *Astrophys. J.*, 1980, **238**, L67 and **241**, L133. The beginnings of lens modelling and the now well known benign angle diagram (figure 2b) for axisymmetric lenses.

Burke, W. L., *Astrophys. J.*, 1981, **244**, L1. The general odd number theorem for smooth lenses. One page, one figure, one reference – a model of elegance!

Young, P. J., *et al., Astrophys. J.*, 1981. **244**, 723.

Narasimha, D., Subramanian, K. and Chitre, S. M., *Mon. Not. R. Astron. Soc.*, 1982, **200**, 941.

Models of the 'triple quasar'. The second paper emphasises the role of weak ellipticity in the potential as the crucial general feature.

Gravitational Lens Optics

Kayser, R and Refsdal, S., *Astron. Astrophys.*, 1983, **128**, 156.

The wavefront point of view for nonsymmetric lenses.

Berry, M. and Upstill, C., *Progress in Optics*, 1980, **18**, 259 (ed. E. Wolf, North Holland) is an excellent review of "Catastrophe Optics" in which the properties of systems without special symmetries are studied.

Arnold, V. I., 1982. Catastrophe Theory (Springer) is a popular (!) book placing rays, wavefronts, caustics and much more in the proper mathematical context of singularity theory.

Blandford, R. D. and Narayan, R., *Astrophys. J.*, 1986, **310**, 568 demonstrate the power of these general ideas in the context of gravitational lenses. Particularly noteworthy is their elegant classification of images via the topology of saddle point contours.

A recent short review of gravitational lens optics is Blandford, R. D., *et al., Science*, 1989, **244**, 824.

Arcs and Rings

Soucail, *et al., Astron. Astrophys.*, 1987, **172**, L14.

Lynds, R. and Petrosian, V., *Bull, Amer. Astron. Soc.*, 1987, **18**, 1014.

These papers describe the discovery of the arcs in clusters of galaxies.

A recent Workshop on Gravitational Lenses (held at Toulouse in 1989, proceedings brought out by Springer) contains accounts of much recent work.