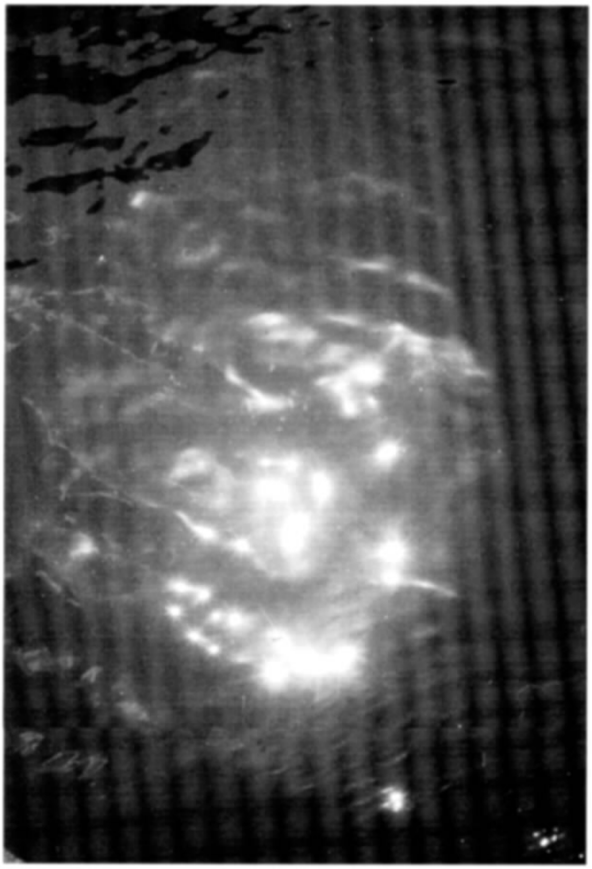
15 Natural focusing

MICHAEL BERRY

I will discuss the geometry of light. In a sense every contribution to this book is concerned with the same subject: eyes are of no use without light, and visual perception inevitably involves geometry. But I am not going to be speaking about light as perceived, as a tool for seeing, but rather concentrating on the physics of light itself. There are several reasons for injecting physics into a book such as this. One is that the patterns I am going to be describing can be seen with the naked eye, unlike much of science nowadays, so it is now 'eyes on', if not 'hands on', physics. Second, the patterns I am describing are appealing ones, although I am not under any delusion that this is either a necessary or a sufficient condition for them to be regarded as art. And third, consistently with this being a section concerned with art and mathematics, the branch of optics I am going to be describing was developed under the stimulus of an advance in mathematics, namely the celebrated and notorious catastrophe theory created several decades ago by René Thom and Vladimir Arnold. What I will describe is the *focusing of light in nature*.

Figure 15.1 shows the sparkling of sunlight on the water in Bristol docks. I am going to use it to sidle into the mathematics. Each of the bright points is an image of the Sun in the water. It corresponds to a place where the water surface has the right slope to direct light from the Sun into the eye. Figure 15.2 shows the sun and the eye in a situation where three rays reach it. If you were to look down you would see three bright points (labelled 1, 2, and 3). What distinguishes these three from all the other points? Of course, you know from elementary optics that rays correspond to points where the ingoing and outgoing rays make equal angles with the reflecting surface. The point P, for example, does not correspond to a ray and has different angles in and out.

But there is another feature which distinguishes the brilliant points. If you ask 'What is the travel time of light between the Sun and the eye via the water, for these different paths?' and draw a graph of it, then you get Fig. 15.3, a landscape curve with maxima



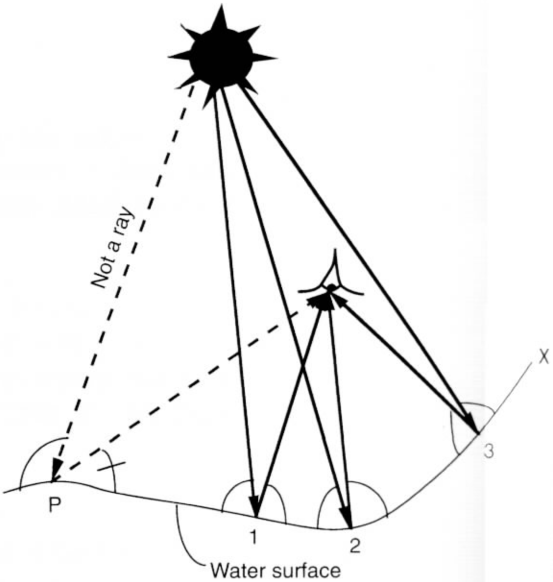


Fig. 15.1. Reflections of the Sun in wavy water.

Fig. 15.2. Three reflected images of the Sun.

and minima. The maxima and minima — places where the landscape is flat, the so-called critical points — are the points on the surface which correspond to rays (this is Fermat's principle of geometrical optics) and so P is not a ray because it is not one of the flats: it lies on a hillside. So there is a relation between optics and the geometry of landscapes. Now, catastrophe theory is about geometric landscapes, so we are in the right mathematical world, at least. But we are not there yet.

To come closer, consider what happens as time passes: the water surface changes and the brilliant points move about. They can collide with each other and disappear or they can be born spontaneously in pairs. These events are called twinkles and they usually happen too fast to be perceived individually. Their rapid succession is what gives illuminated water its sparkling appearance. Now, each twinkle corresponds to something more than the water surface having the right slope. It corresponds to the water surface also being curved just right, so as to *focus* the light into the eye as well as directing it there.

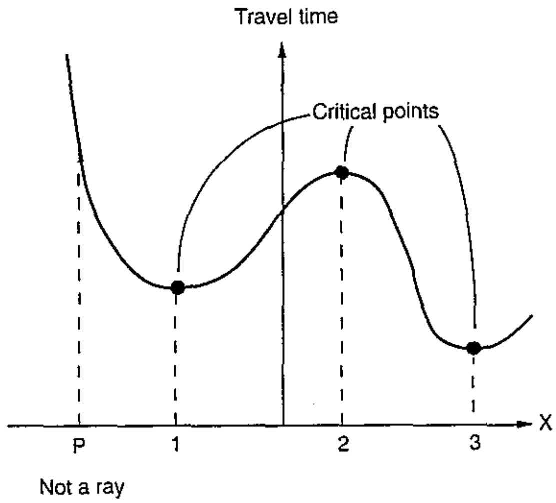


Fig. 15.3. Travel time 'landscape': rays correspond to critical points (minima or maxima).

To illustrate this, Fig. 15.4 shows a different view of the geometry, with the sun far away. Rays emerge in different directions. Their envelope is a focal curve (in three dimensions it would be a surface), called a *caustic* from the Greek work for burning, because curves like this are commonly seen in light focused by burning glasses. The caustic distinguishes places like A, where only one reflected ray exists (one brilliant sun image) from places like B, where there are three. Figure 15.5 is a water surface seen from a place like B. This has local interest, because it shows the tidal bore on the River Severn near Bristol, together with a TV helicopter and its three reflections (the third is very faint). Across the caustic there is a discontinuous change in the number of rays. Commonly the caustic is more complicated and one sees many images. So the existence of caustic surfaces in space lies at the heart of the births and deaths of the brilliant points.

We shall concentrate on this kind of focusing. Let us look more closely at one of the caustic curves (Fig. 15.6) and see what happens near it. The caustic is the envelope of the family of rays. Across it, the number of rays changes by two. The caustic is very bright, because the energy that is concentrated between any pair of rays shrinks down to be infinitely concentrated there. Therefore, caustics dominate optical patterns, as we will see. They are places of discontinuity and, moreover, of exactly the same discontinuities that the mathematics of catastrophe theory describes, namely discontinuities in mathematical landscapes as they are altered to make maxima and minima coincide.

Therefore the mathematical theory of catastrophes describes the physics of focusing. What does it tell us? It gives us a classification of the geometric shapes that caustics can take. There is one refinement

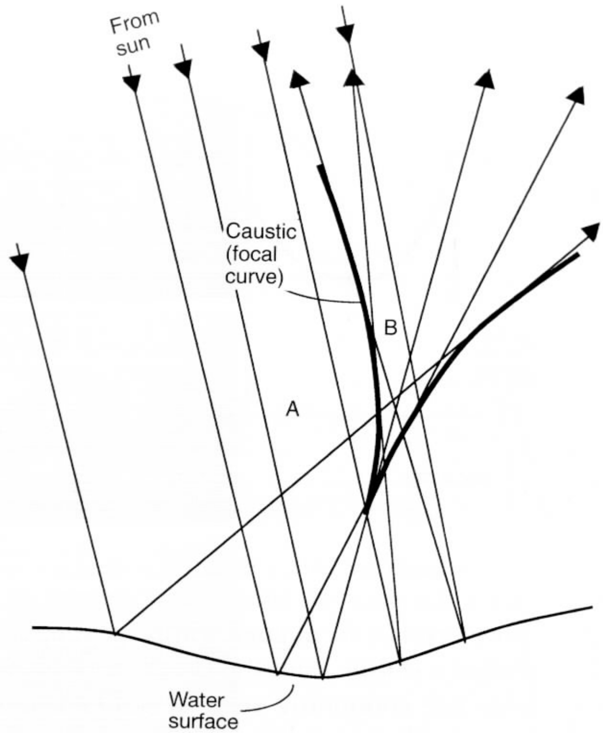


Fig. 15.4. Caustic curve of focused reflected sunlight.

Fig. 15.5. Reflections in the Severn Bore.

which is actually crucially important. Instead of classifying all focal surfaces, it classifies those which have a mathematical property called 'structural stability'. It is often easy to understand something by knowing what it is not. Figure 15.7 shows an *unstable* focus, the

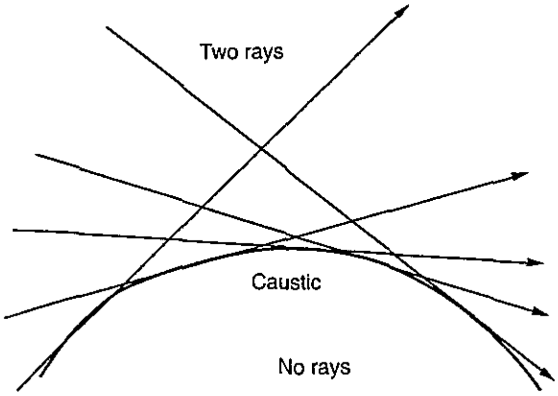


Fig. 15.6. Across a caustic, the number of rays jumps by two.

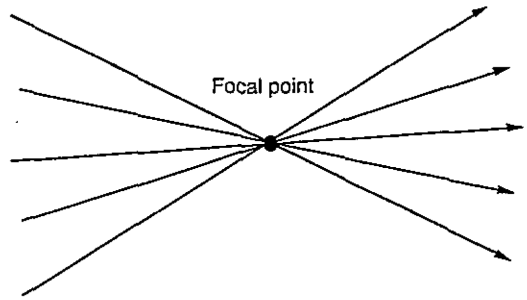


Fig. 15.7. Unstable point focus.

one we learned about in school optics, where all the light rays pass through a single point. That is highly improbable in a focusing situation; a large part of the lensmaker's art is devoted to making it happen. Structural instability means that the slightest disturbance of the conditions that gave rise to it caused this point to explode into complicated shapes, which are what the theory classifies. So we expect that the shapes that the mathematics classifies are those caustic curves and surfaces to be found in nature, where there are often no special symmetries.

Let us start to go through the catalogue. You might expect the mathematical caustics to be organized in terms of the number of dimensions they occupy. That is not quite how it works. They are organized by something called *codimension*. The codimension is the number of dimensions you must typically explore to find something. For example, a line in space has codimension two, because in order to cut it you typically need a surface, which has two dimensions. Obviously the simplest case is codimension one. The corresponding catastrophe (Fig. 15.8) has a name: the *fold*. A point on a line has codimension one. You find it as you go along a line in one dimension. A line in a plane has codimension one. You find it by exploring the plane. Similarly, a surface in space has codimension one. This is not very exciting geometry; there seems nothing particularly interesting about it. In focusing, however, it corresponds to something very familiar, namely the *rainbow*. If anybody asks you what catastrophe theory is, you can say it is about rainbows. I will explain why that is, before going on to more substantial applications.

The rainbow is formed by sunshine on a dripping cloud. Sunlight hits a raindrop at various latitudes and comes out in various directions. If you make a graph of the out direction against the in latitude then you find a curve with a fold (minimum) in it. Near the fold a bundle of rays going in gets concentrated into a very narrow bundle going out. This is focusing in angle. The angle is a single variable

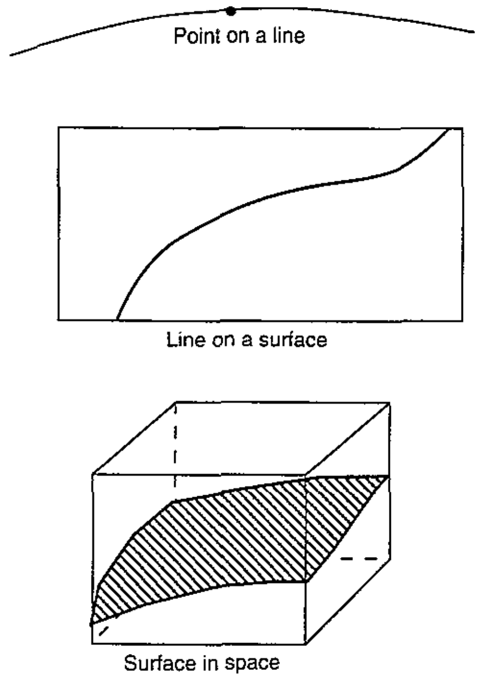


Fig. 15.8. Fold catastrophe: co-dimension one.

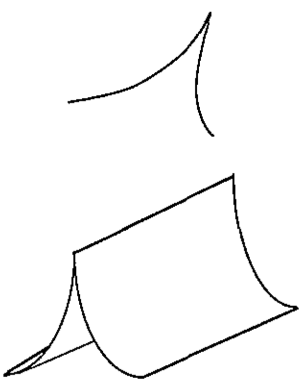


Fig. 15.9. Cusp catastrophe: codimension two.

that you change in order to see the caustic, which is at 138° to the forward direction, that is away from the sun. The raindrop reflects a bright cone of rays at 42° to the backward direction, and you, on the ground, see, brightly lit, all the raindrops on whose cones your eyes lie.

Now let us go on to codimension two. This is important because screens, retinas, and photographic plates are two-dimensional. Here the catastrophe is called the *cusp* (Fig. 15.9). In the plane, a cusp is a point where two fold caustic curves meet: they kiss, with a common tangent. In three dimensions this codimension-two object is a cusped line, a crease where two focal surfaces join. Here the mathematics gains substance, because it declares that this cusp is the only stable structure: other structures are not stable. For example, a smooth curve coming to an end (like the crock of gold at the rainbow's end) is not stable, nor is an isolated point, nor is a finite-angled corner: these cannot exist unless there is some special condition operating. It is quite easy to see cusps. You can see them as bum-shaped curves in teacups (Fig. 15.10), diffusely reflected from the milk in the tea or coffee. You can also see them if you wear glasses and walk in the rain at night and look through the distorted droplets on your glasses (which were not as clean as you thought, so that — because of non-uniform wetting — the drops are not circular). Look at a distant light through such water-droplet lenses and you will see patterns like Fig. 15.11 with many cusps. In Fellini's film $8\frac{1}{2}$, there is a beautiful sequence at night where

Fig. 15.10. Caustic in a teacup.

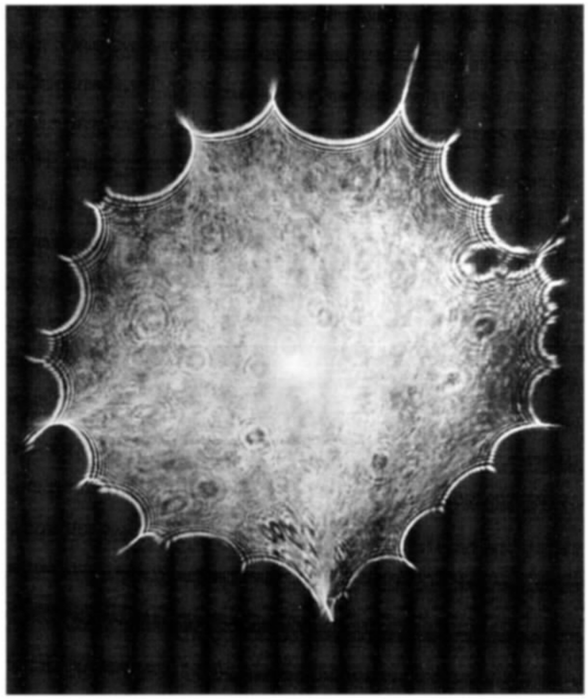


Fig. 15.11. Cusps from an irregular water-drop 'lens'.

the camera in the rain is looking at circus lights; cusps abound. Clearly visible in Fig. 15.11 are delicate interference bands. These will be more prominent in later pictures and I will describe them later. At the moment, however, I am drawing attention to the cusps, as examples of typical focusing. You can create cusps easily with a laser beam shining through a water-droplet lens hanging on a dusty glass plate (that is how Fig. 15.11 was made).

Let us proceed to three dimensions. There are three different co-dimension-three catastrophes; that is, three different types of singular point on caustic surfaces in space. One is the *swallowtail*, so named by the blind mathematician Bernard Morin. In the swallowtail (Fig. 15.12) two cusped edges meet at the catastrophe point. It is hard to show caustics in three dimensions. It is possible to see them with smoke in the region of the focus, but much more commonly we observe sections and the telltale section of the swallowtail is a self-crossing of caustic curves. Again, it is possible to make swallowtails with water droplets; Fig. 15.13 shows two.

The second codimension-three catastrophe is the *elliptic umbilic* (Fig. 15.14), a waisted spire shape, whose telltale sections are three-cusped triangles. There is a special section which is just a point. Previously I stated that a point focus is unstable and now we can see the instability in action: with any other section, the point explodes into the three-cusped triangle, which is stable because it consists

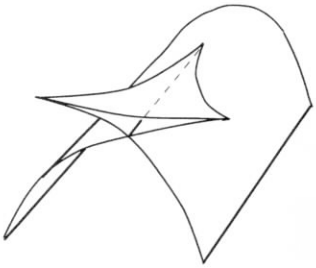


Fig. 15.12. Swallowtail catastrophe: codimension three.

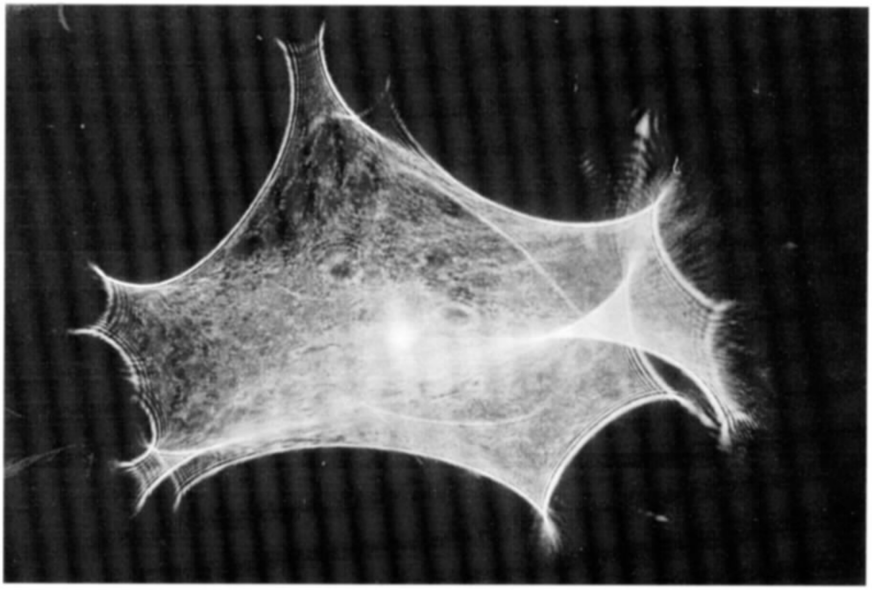


Fig. 15.13. Swallowtails from an irregular water-drop 'lens'.

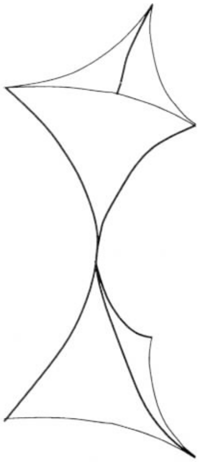


Fig. 15.14. Elliptic umbilic catastrophe: codimension three.

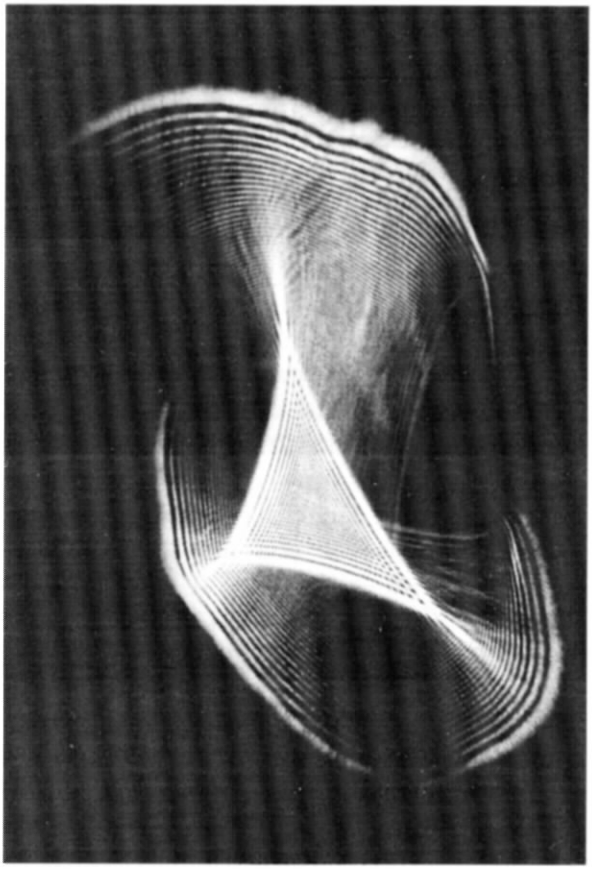


Fig. 15.15. Elliptic and hyperbolic umbilics from bathroom-window glass.

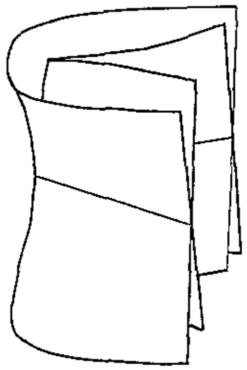


Fig. 15.16. Hyperbolic umbilic catastrophe: codimension three.

only of folds and cusps Figure 15.15 shows an elliptic umbilic section, produced by shining light through smoothly irregular bathroom-window glass. The three-cusped triangle is obvious, as is much interference decoration.

Figure 15.15 also serves to illustrate the third and last co-dimension-three catastrophe, namely the *hyperbolic umbilic* (Fig. 15.16), which consists of two interpenetrating sheets, the inner of which is cusped. So the telltale section here is a smooth outer curve and a cusped inner one. There is also a special (unstable) section with a finite-angled corner, where the two curves coincide. Figure 15.15 shows three hyperbolic umbilics, in various degrees of unfolding from their special section, sprouting from the elliptic umbilic triangle.

This kind of physics has a botanical flavour: the mathematics guarantees that certain shapes are stable and one does experiments with no special symmetry operating and seeks specimens of them. Different 'specimens' of the same catastrophe, for example the three hyperbolic umbilics in Fig. 15.15, are the same in the sense that flowers of a particular type are the same, and not in the sense that Ford cars or crystals are the same. What the catastrophes share is a topological identity with a precise mathematical meaning (basically, any 'specimen' can be obtained by smoothly deforming any other specimen of the same catastrophe).

It is not at all far-fetched and indeed quite helpful to think of the catastrophes as *atoms of form*. Consider real atoms. They occupy a sort of mesoscale, separating larger (macro) scales, on which atoms link together into molecules, liquids, solids, etc., from smaller (micro) scales, with subatomic detail such as electrons, nuclei, quarks, etc. Likewise in optics the catastrophes occupy a mesoscale. As we will see later, many catastrophes can be linked together into larger structures. Figure 15.15 shows a rudimentary example, in the form of a 'molecule' consisting of one elliptic and three hyperbolic umbilics. It is also possible to descend to finer scales, and this we now do.

Until now, although the mathematics I have been describing is today's, the physics has been that of the seventeenth century, namely the ray theory of light. Now we move on to the physics of 1800, when Thomas Young, in experiments at the Royal Institution, just off Piccadilly, proved that light is better described as a wave motion, with a wavelength of about one two-millionth of a metre. Of course we have already seen these waves as fine-scale decorations of the caustic patterns in Figs. 15.11, 15.13, and 15.15. One of the pleasant surprises from catastrophe theory is that the mathematics gives not only a classification of the caustics, but also a description of the wave interference that decorates them on fine scales: each caustic has its characteristic interference pattern.

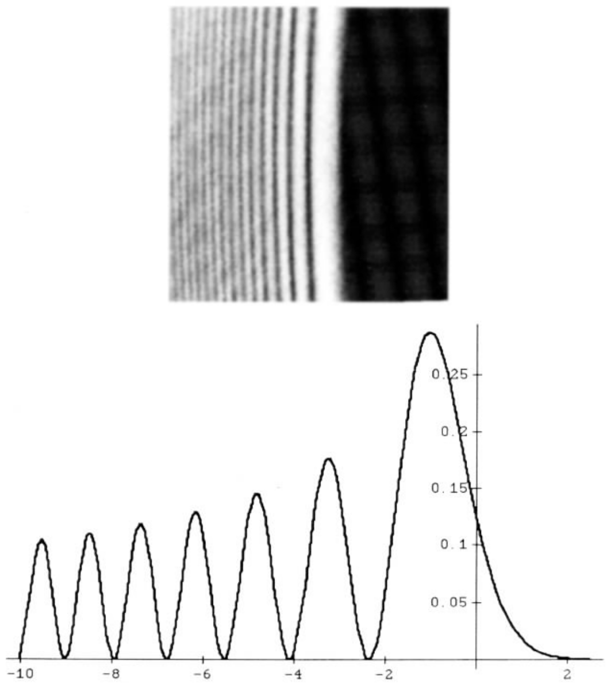


Fig. 15.17. Interference fringes near a fold catastrophe and graph of light intensity.

Figure 15.17 shows the pattern that decorates the fold catastrophe. If ray theory were true, the intensity would rise smoothly to the bright line, where geometric focusing would occur, but in reality there are interference fringes on the side where there are two rays. In the case of the cusp (Fig. 15.18(a)), the pattern is much more complicated. Magnification (Fig. 15.18(b)) and comparison with a computer simulation generated by the mathematics (Fig. 15.18(c)) shows that the smallest details, even the little pairs of dark spots, which are only a few wavelengths apart, can be accurately reproduced. One is here reminded of Maxwell's observation: 'the dimmed outlines of phenomenal things all merge into one another unless we put on the focusing glass of theory and screw it up, sometimes to one pitch of definition and sometimes to another, so as to see down into different depths through the great millstone of the world'.

In codimension three, the interference catastrophe patterns are incredibly complicated architectural arrays of brights and darks in space. We can only look at sections. Figure 15.9 shows part of the elliptic umbilic pattern, alongside a computer simulation. The whole three-dimensional pattern has been studied in great detail, as have its counterparts for the hyperbolic umbilic and the swallowtail.

Although this will mean digressing from our main theme of focusing, I cannot resist bringing in the physics that we have learned since 1800. We now know from quantum mechanics that light is

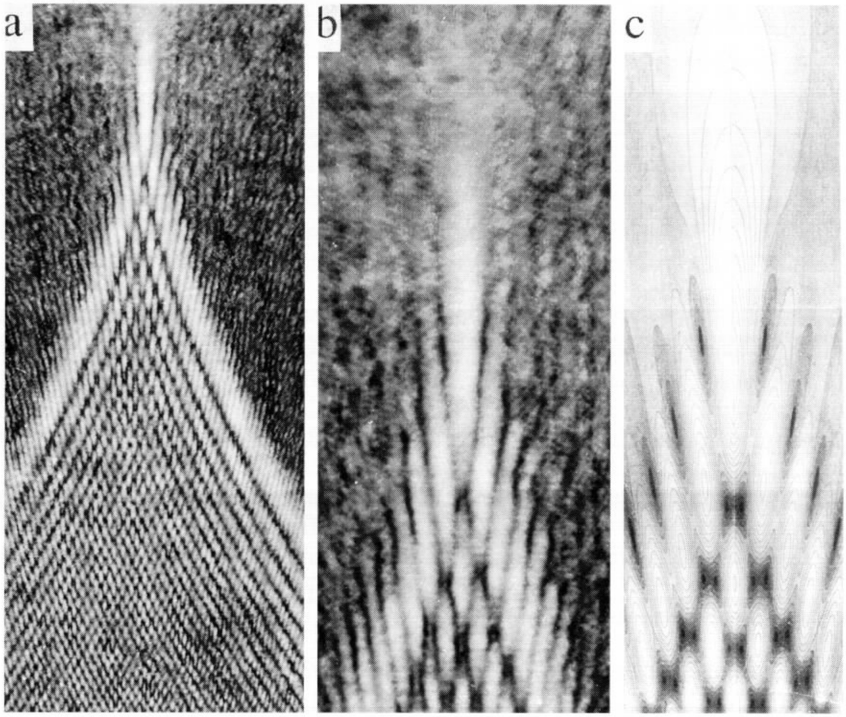


Fig. 15.18. Interference pattern near a cusp catastrophe.

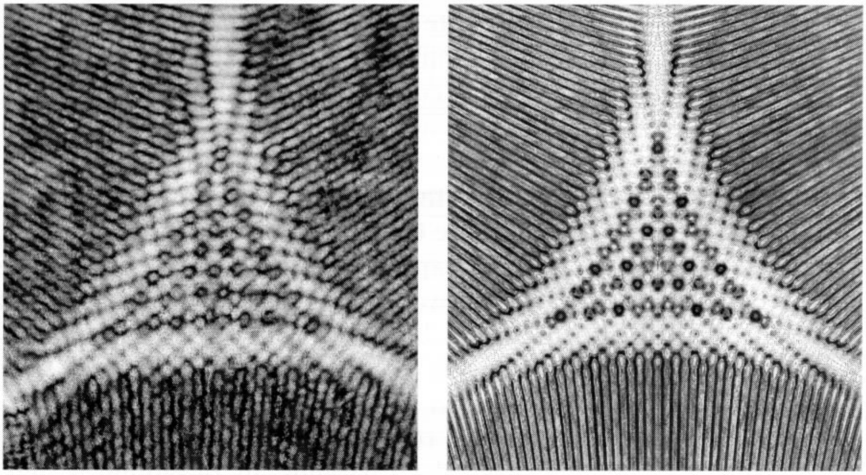


Fig. 15.19. Section through the interference pattern near an elliptic umbilic catastrophe (left); computer simulation (right).

actually a stream of photons. Why can these not be seen in the fine-scale interference patterns? The reason is that the light is too bright. In the laser beam that produced Fig. 15.17, about 10 000 million million photons strike each square millimetre of the screen every second, so we see only their collective effect. To detect individual photons, it is necessary to dim the light by a comparable factor. Then the interference fringes would be built up point by point. Photons

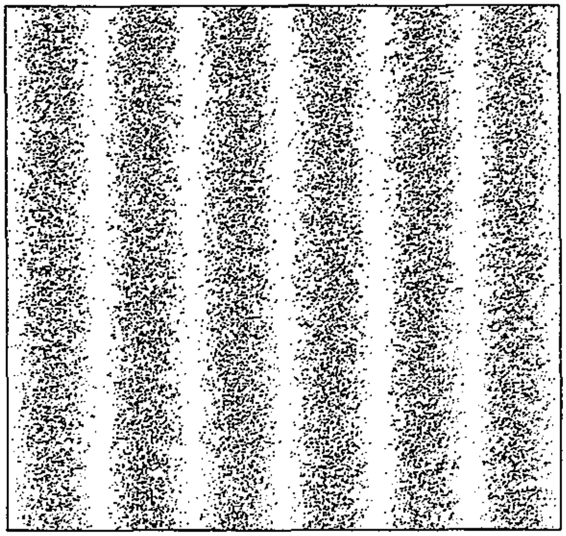


Fig. 15.20. Interference fringes built up by random impacts of photons.

arrive at random, but after some time it becomes apparent (Fig. 15.20) that the probability of arriving near a bright line is much greater than that of hitting a dark line. So the pattern you see is actually built up out of a multitude of individually random events.

To finish, let us climb to the macroscale and look briefly at the intricate networks of caustics produced on bottoms of swimming pools on sunny days. The sunlight refracted by the water surface makes complicated caustic surfaces in the space in the water below and what we see is where those patterns cut the bottom of the pool, which acts as a screen. Rather than show a real caustic network, it is instructive to begin with David Hockney's painting of one (Plate 17), because this illustrates an important point.

As expected, no interference detail is visible on such a large scale. What was not expected, however, is that the patterns as observed cannot be immediately interpreted as stable caustics. Consider, for example, the many places in Plate 17 where caustic lines appear to meet in threes. Such meetings are not described by catastrophe theory (the only stable singularities in the plane are cusps). This looks like a paradox. Its resolution is that we usually see swimming-pool caustics under conditions of poor resolution, which obscure their fine structure. The blurring is caused by the rapid motion of the patterns and the half-degree width of the Sun's disk.

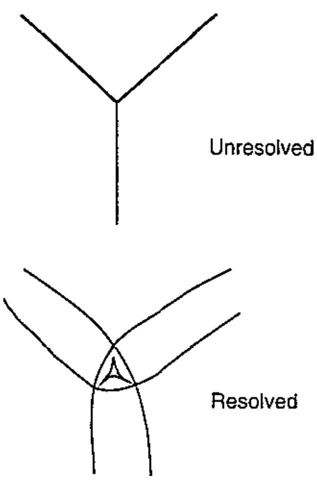


Fig. 15.21. Caustic 'triple' junction, with elliptic umbilic at its centre.

More detailed laboratory investigation of this caustic junction, supported by mathematics, reveals that each of the lines is really double, so six of them meet where uneducated observation suggests three. At each meeting-point there is an elliptic umbilic triangle (Fig. 15.21). There are other sorts of junction, and complicated ways in which they can be assembled into networks. In every case, the large-scale networks are organized by the forms in the small library of catastrophes, acting as structural elements enabling us to understand what would otherwise be collections of lines without meaning.

Further reading

Arnold, V. I. (1986). *Catastrophe theory* (2nd edn). Springer, Berlin.

Berry, M. V. (1976). Waves and Thom's theorem. *Advances in Physics*, **25**, 1–26.

Berry, M. V. and Nye, J. F. (1977). Fine structure in caustic junctions. *Nature*, **267**, 34–6.

Berry, M. V. and Nye, J. F., and Wright, F. J. (1979). The elliptic umbilic diffraction catastrophe. *Philosophical Transactions of the Royal Society of London*, **A291**, 453–84.

Berry, M. V. and Upstill, C. (1980). Catastrophe optics: morphologies of caustics and their diffraction patterns. *Progress in Optics*, **18**, 257–346.

Nye, J. F. (1978). Optical caustics in the near field from liquid drops. *Proceedings of the Royal Society of London*, **A361**, 24–41.

Nye, J. F. (1986). The catastrophe optics of liquid drop lenses. *Proceedings of the Royal Society of London*, **A403**, 1–26.

Poston, T. and Stewart, I. (1978). *Catastrophe theory and its applications*. Pitman, London.

Walker, J. (1983). Caustics: mathematical curves generated by light shined through rippled plastic. *Scientific American*, **249**, 146–53.

Walker, J. (1988). Shadows cast on the bottom of a pool are not like other shadows. Why? *Scientific American*, **259**, 86–9.

Walker, J. (1989). A drop of water becomes a gateway into the world of catastrophe optics. *Scientific American*, **261**, 120D–3.

Upstill, C. (1979). Light caustics from rippling water. *Proceedings of the Royal Society of London*, **A365**, 95–104.

Wright, F. J. (1988). Spectacles in the rain: catastrophe optics. *Physics Bulletin*, **39**, 313–16.