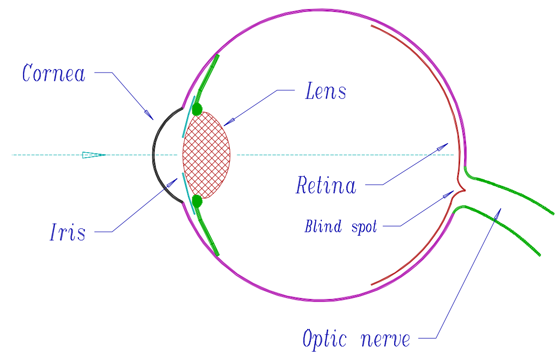
The human visual system

In the course of several billion years, Earth's life forms have evolved numerous schemes for using electromagnetic radiation in the wavelength range 3500-9000 Å (infra-red through ultra-violet light). Plant chlorophyll, bacterial rhodopsin, photosensitive spots, compound insect eyes, camera-like eyes, and the pits of a pit viper----all exist to transform light into something else.

For plants and bacteria, light provides the energy that powers their chemical factories, through the process known as photosynthesis[1]. In higher organisms, eyes and their equivalents transduce light into nerve impulses, thereby enabling the possessors of a visual sense rapidly to detect changes in their environment.

Such changes can presage danger, indicate a source of food, a mate, resynchronize the circadian clocks----all of which are intimately connected with survival. Therefore vision must have been under intense evolutionary pressure to improve to an optimum level consonant with the overall economy of the organism. That is, Darwin's concept of the origin and improvement of vision through natural selection explains why the eyes of higher organisms perform so well, without the need of teleological argument.

In this chapter we study the mammalian eye----primarily the human eye----explaining its functioning in terms of optics and quantum mechanics. We shall be particularly interested in discussing the various optimizations and engineering compromises to be found in eyes.



Horizontal section of the right eye, from above

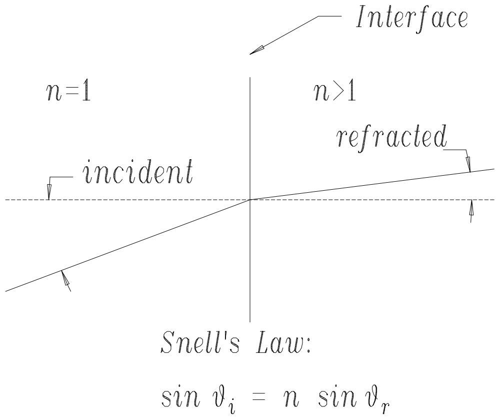
1. Simple lenses

As the sketch to the above right shows, the human eye is constructed along much the same lines as a box camera: light entering through a hole at the front (the pupil) is focused on a light-sensitive layer at the back (the retina). The simplest approach to the physics of lenses treats light as *rays* rather than electromagnetic waves. That is, we imagine that light emanating from a point moves through a homogeneous transparent medium in a straight line.

The key to understanding lenses is what happens when light falls on an interface between different media. To begin with we imagine the interface is a plane surface, as shown on p. 100. *Snell's Law* relates the angle between the normal to the surface (dashed line) and the incident ray, to the angle between the normal and the refracted ray:

$$1 \sin\theta_i = n \sin\theta_r ,$$

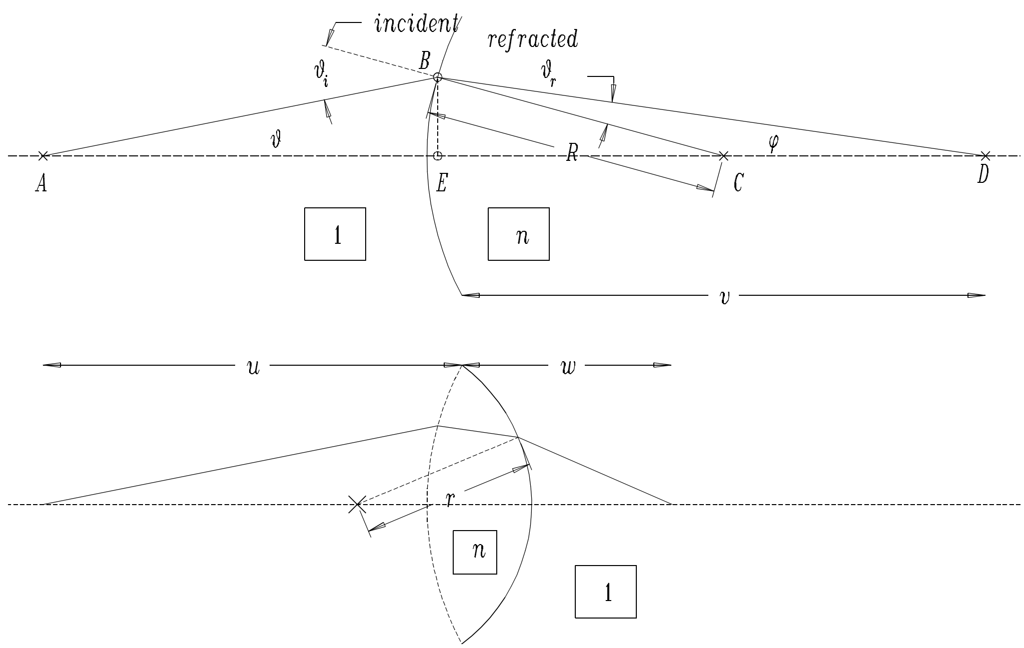
1. You might not be aware that humans also employ photosynthesis: under the influence of photosynthesis, made possible by ultraviolet rays from the Sun, a sterol compound from the liver (dehydrocholesterol) is converted to vitamin D3. This supplies enough vitamin D3 for human needs.



where n is the index of refraction of the medium and 1 is the index of refraction of the vacuum.

The keys to understanding lenses (that is, refraction at curved interfaces) are first, to regard small patches of the surface as flat; and second, to treat the two curved surfaces separately.

Consider the upper of the two figures below. We



need to relate the distances u and v to the radius R of the lens surface (which for simplicity we take to be spherical) and the index of refraction n. The triangle ABC has sides $u/\sin\theta$, $u+R$ and R, respectively. From the law of sines we have

$$(u+R)\,\sin\theta \;=\; R\,\sin(\pi-\theta_i) \;\equiv\; R\,\sin\theta_i\,.$$

The triangle BCD gives us the relation

$$R\,\sin\theta_r \;=\; (v-R)\,\sin\varphi\,.$$

We also see that the shared altitude of the two right triangles, ABE and BDE, is

$$\overline{BE} \;\equiv\; \overline{AB}\,\sin\theta \;=\; \overline{BD}\,\sin\varphi\,.$$

Hence in the limit of small angles

$$\frac{u}{v} \approx \frac{\sin\theta}{\sin\varphi}$$

or

$$\frac{u+R}{v-R} \approx n\,\frac{u}{v}$$

leading to the relation

$$\frac{1}{u} + \frac{n}{v} \approx \frac{n-1}{R}\,.$$

Now to combine the effects for two curved surfaces we simply apply the formula again, but this time keeping in mind that for the right-hand surface the image distance and radius of curvature must be considered negative: that is, $w < 0$ and $r < 0$, whence

$$\frac{n}{v} - \frac{1}{|w|} = \frac{n-1}{-|r|}\,.$$

When we combine the two we get

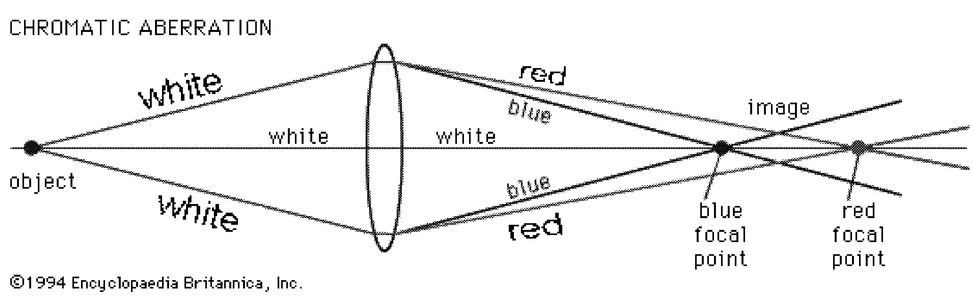
$$\frac{1}{u} + \frac{1}{|w|} = (n-1)\left(\frac{1}{R} + \frac{1}{|r|}\right) \overset{df}{=} \frac{1}{f}$$

where f is called the *focal length* of the lens.

The physical interpretation of the focal length is it is that point where light rays coming from infinitely far away ($u = \infty$, that is parallel rays) are brought to a point; or alternatively if a point source of light is placed at the focal distance on one side of the lens, parallel rays emerge on the other side.

2. Problems with lenses

A simple lens made of a uniform material, whose surfaces are elements of a sphere, suffers from several problems in forming images. The easiest to understand is *chromatic aberration*: in general the index of refraction of a transparent material depends on the wavelength of the light. The focal length is different for each color, hence the lens focuses multicolored light to a rainbow-like ring (the best compromise) rather than to a point. This reduces the resolution. High quality camera lenses compensate chromatic aberration by joining layers of different kinds of glass, with opposite variations of refractive index with wavelength. The opossite effects tend to cancel.



A second difficulty is *spherical aberration*---light rays passing through the lens far from its axis converge at a different point from axial rays. To some extent this can be compensated by restricting the entrance aperture of the lens so that rays pass through close to the axis. (It can also be reduced by varying the shape of the lens from spherical.)

A third proble is *astigmatism*---rays coming from points off the axis are not imaged at the same distance behind the lens as rays originating near the axis. This effect can be compensated to some extent by choosing the radii of the two surfaces carefully, by restricting the lens aperture, and also by varying the lens shape from spherical.

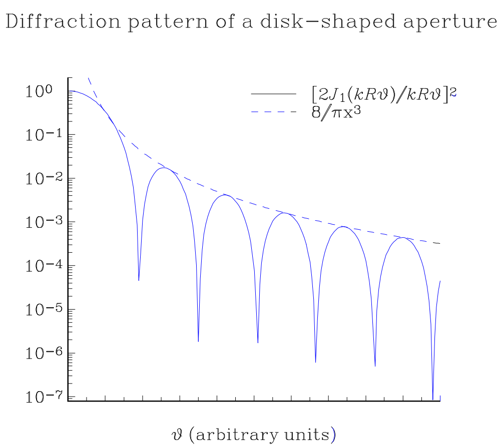
Not all the light incident on an interface between media of different indices of refraction gets transmitted. Some is reflected at each surface. The bigger the didiscontinuity of refractive index, the more light is reflected. Since the best lens materials have high refractive index, they tend to reflect the most light. Thus lenses suffer from loss of light intensity, excluding whatever light is absorbed by the medium (since no medium but vacuum is perfectly transparent). Manufacturers of high quality lenses apply various surface coatings to lessen the interface discontinuity. By properly adjusting the spatial variation of n in the transition region, such coatings can reduce substantially the reflectivity of the surface[2].

How important are these innate defects in optical systems? In cataloguing them we have neglected one important limitation on the resolving power of

2. This is rather like the special paint that is non-reflective at radar frequencies used on “stealth” aircraft.

lenses, the wave character of light. As Lord Rayleigh showed, light passing through a circular aperture produces a diffraction pattern of concentric rings described by the intensity function

$$I(\theta) = I_0 \left(\frac{2\,J_1(kR\theta)}{kR\theta} \right)^2$$



where the "wave number" is $k = \frac{2\pi}{\lambda}$. This function is graphed below.

Rayleigh noted that light from a point source will image to a diffraction pattern because of the finite size of the lens. The larger the radius, for a given wavelength of light, the smaller will be the angular size of the central maximum. If two sources are so close that the central maxima of their images overlap, then clearly they cannot be distinguished. Since the first zero of the Bessel function $J_1(x)$ occurs at $x = 3.8317\ldots$ the angular size of the centralmaximumis

$$\theta_{min} = \frac{3.8317\ldots}{2\pi}\,\frac{\lambda}{R} \approx 1.22\,\frac{\lambda}{D}$$

where *D* is the lens diameter.

For the human eye the angular size of the central maximum from a pointlike source is minimum when the pupil is at maximum size, *i.e.* in dim light

$$\theta_{min} \approx 1.22 \times \frac{5\times10^{-5}\text{ cm}}{1.0\text{ cm}} = 6.1\times10^{-5}\,.$$

Since the eye is about 2 cm from lens to retina, the spacing between neighboring cone cells on the retina (at the *fovea*, or point of greatest visual acuity—see Figure on p. 99) should be about 1.2×10^{-4} cm apart—closer spacing would not improve the eye's resolution. The diffraction criterion thus predicts a cell density of 1.7×10^{7} per cm^2.

Now suppose the various reductions of resolving power mentioned previously—chromatic[3] and spherical aberration, and astigmatism—were so severe as to overwhelm the Rayleigh diffractive limit. In that case we should expect a much lower density of receptors than $\approx 10^7/\text{cm}^2$. However, the receptor density in the fovea is $1.6\times10^7/\text{cm}^2$; hence we may surmise[4] that diffraction is the limiting factor. In other words, Nature's engineering here—as in every other aspect of physiology we have studied—is only as good as it needs to be.

3. Problems with vision

Some visual defects are optical in origin and may be corrected with external lenses or prisms, contact lenses, or—lately—corneal surgery. *Myopia* (near-sightedness) occurs when the eyeball is too long from lens to retina, so that the focal point for distant objects is in front of the retina. Nearby objects, however, can be focussed correctly at the retina. *Hyperopia* (far-sightedness) arises from a lens that is too close to the retina, so that distant

3. Actually, the fovea is yellow in color—that is, it has a filter that restricts the range of frequencies it accepts. We presume this adaptation inhibits color aberration!

4. …even without attempting to measure directly the other types of aberration…

objects are focussed properly but near ones are fuzzy. The simple lens equation,

$$\frac{1}{u} + \frac{1}{w} = \frac{1}{f},$$

makes the effects of these defects obvious, and also makes obvious how to correct them: the near-sighted person places a diverging lens (plano-concave or bi-concave) before his eye so that distant light is focussed at the retina. The far-sighted person requires a converging lens (plano-convex or bi-convex) to do the opposite.

In middle age, the lenses tend to lose their flexibility, and the eye muscles can no longer make them change their focus. They generally stiffen in the "far-sighted" conformation, so that even formerly myopic persons begin to be able to see distant objects. But they can no longer image nearby objects and find reading difficult or impossible. This condition is called *presbyopia*.

The invention of corrective lenses (in Italy, near the end of the 13th Century[5]) was an enormous boon to literate persons who formerly had to employ young secretaries to read to them, once they reached middle age. Presumably the fact that many of the literati were churchmen accounts for the fact that optometry was not considered witchcraft.

A third kind of visual defect is *astigmatism*. Either the cornea or the outer surface of the lens may be non-uniformly curved, so that images will be distorted. Mild astigmatisms tend to be compensated by the brain—that is, the image is "computer-corrected". Severe cases require specially shaped compensating lenses, or sometimes laser surgery to reshape the cornea.

There is not space to discuss visual problems that arise from damage to the retina or optic nerve, such as macular degeneration, glaucoma, or diabetic retinopathy. The subsequent section will make clear why such disorders cause blindness.

We also must, for the most part, bypass the fascinating subject of color vision and color blindness.

4. Light → nerve impulses

The human retina is a complex structure. The outermost layer of retinal cells contain the pigment melanin[6] so that no light is reflected from the back of the eye (reflected light—as in the eyes of nocturnal predators such as cats and sharks—would produce double images that might confuse our vision). The next layer are light receptors. Then come several layers of neurons that perform various kinds of pre-processing before the visual signal is sent down the optic nerve to the lateral geniculate nuclei in the brain.

Curiously, our neurons lie between the incoming light and the photoreceptor cells. That is, the light-sensitive cells are at the back of the eye. It does not have to be this way: in the octopus, whose eyes are startlingly similar to human eyes, the photocells precede the neurons in the optical path. We realize from this that human and octopus eyes exemplify parallel evolution—we are not descended from mollusks!—and their similarities result from the constraints imposed by physical law.

5. Fra Giordano of Pisa said, in a sermon dated 1306, "It is not twenty years since there was found the art of making eyeglasses which make for good vision, one of the best arts and most necessary that the world has. So short a time it is since there was invented a new art that never existed. I have seen the man who first invented and created it, and I have talked to him."
6. That this pigment is necessary is shown by the fact that albinos—who lack it—have difficulty seeing under certain light conditions that normal eyes handle satisfactorily.

The primary light receptors are of two kinds: cone cells distributed mainly at the fovea (see Figure on p. 99), and rod cells distributed at larger angles off the axis of the eye. The names of these cells reflect their appearance under the microscope---cone- or rod-shaped.

The rod cells are the more sensitive to light, hence are our primary means of seeing in dim light. However they do not discriminate with regard to wavelength. Our color vision comes entirely from the cone cells. Because the rod cells are found off the central axis of the eye, we see best in poor light conditions through our peripheral vision.

The key experiment that first measured the quantal nature of light detection was reported in 1942 by Hecht *et al*[7]. Among other facts they elucidated about the eye, they discovered that for light flashes lasting less than about 0.1 sec, the eye measures only the total light in the flash (that is, the total energy). Their experiment consisted of allowing light flashes of measured total energy to fall off-axis, on rod cells of the experimental subject's eye. The subjects were given a long period in the dark so their eyes would become dark-adapted.

The efficiency of the eye in harvesting photons is only about 10%. That is, only about 50% of the photons incident on the eye make it to the retina (the rest being reflected at the cornea or lens, or absorbed in the vitreous humor). Of the remainder, only 20% are converted into nervous impulses by the receptor cells. Considering the comparable efficiency of the best solid state light detectors, 20% efficiency is actually quite remarkable.

The authors then determined the probability of a subject seeing a flash, as a function of the intensity of the flash. Plotted on a logarithmic horizontal scale, the curve looks something like those plotted to the right. We can understand the probability of seeing in terms of the Poisson distribution. Suppose the eye has a threshold of m_0 photons in order to register a response to light. Now when we think we allow a number m of photons into the eye, this is actually a Poisson-distributed process---really we should say we have arranged matters to allow

$$\overline{m} = \frac{1}{h\nu} \int_{\Delta t} dt\, I(t)$$

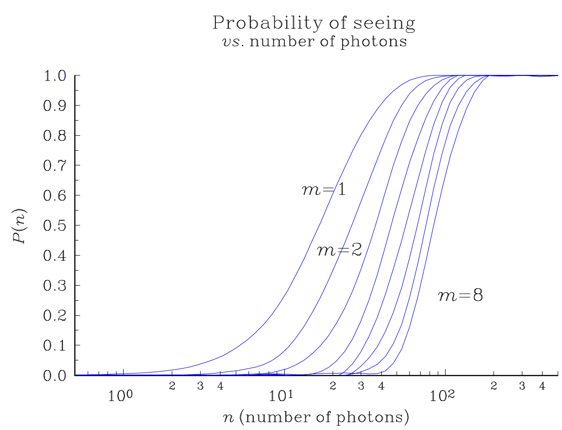
photons into the eye, where $h\nu$ is the energy per photon and Δt is the time the shutter is open. Then the probability that m actually enter is

$$p_m = (\overline{m})^m \frac{e^{-\overline{m}}}{m!} .$$

Hence the probability that more than the threshold number enter is

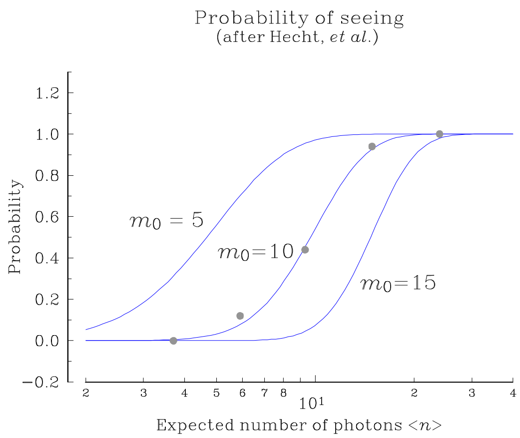
$$P(m \geq m_0) = \sum_{m_0}^{\infty} p_m .$$

These are the curves shown below, with different values of m_0. The best fit to the measured curve



of Hecht, *et al*. was obtained with a threshold value between 5 and 11 (overleaf). They concluded that about 7 photons must be absorbed in order to produce a sensation of seeing.

7. S. Hecht, S. Shlaer and M. Pirenne, "Energy quanta and vision", *Journal of General Physiology* **25** (1942) 819.



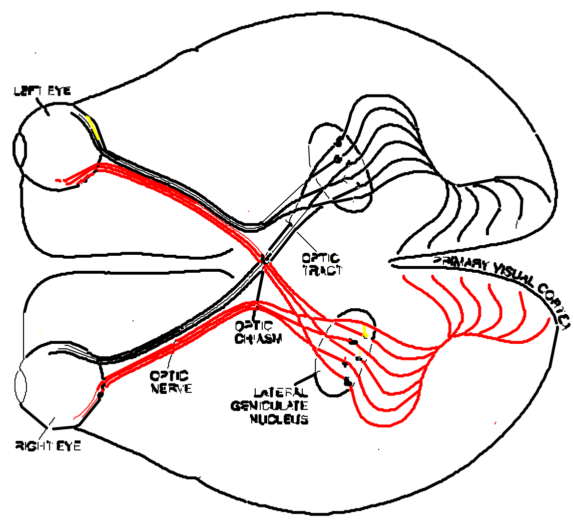
5. Vision and the brain

Visual information must be processed by the brain before it can be of any use to a living organism. There are some 6 million cone cells, but only about one million ganglia that collect their information. Let us suppose that the optic nerve must encode about 2 million bits of information to present a color visual image. With a response time of 0.1 seconds, this means the eye is encoding information at the rate of 20 megabits/sec. Since the switching time for neurons is no faster than 0.01 sec, this implies that at least several hundred thousand neurons are performing the processing in parallel.

In addition to purely visual information, the eyes also convert angular information into distance estimates. That is, the distant rhinoceros is far less alarming than one only twenty yards away.

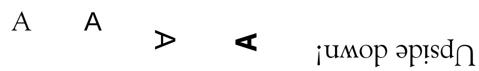
One way that parallel processing can be performed in real time[8] is by mapping the signals from contiguous cells in the retina onto a spatially striated region of the brain's visual cortex, as shown to the right. Moreover, because binocular vision was evidently crucial to our arboreal ancestors (who had to judge distances accurately in order to swing through trees) and equally crucial to our subsequent lifestyles, the optic nerves split, as shown to the right. Half the information from one eye is combined with half that from the other in bodies called lateral geniculate nuclei. Once the processing needed to combine the images into one has been accomplished therein, the visual signals continue to the visual cortex at the back of the brain.

What is the purpose of the spatially striated pattern that is imposed on the visual signals by this pattern of interconnections? We may reason by analogy with the action of the cochlea, that analyzes sound into its component frequencies by spatially differentiating the frequency response of the basilar membrane. Possibly the pattern of striations is imposed on the visual field to perform a spatial Fourier transform, suitable for locating patterns. The ability of the brain to recognize patterns such



8. In the language of embedded computing, "real time" processing has to take place fast enough to be useful. That is, it is useless to discover that a lion is approaching after one has been eaten.

as letters of different forms and in different orientations

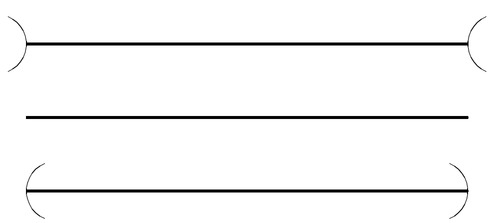


is something we do not at all understand, and have so far been unable to duplicate in computing machinery. But our eyes and brain accomplish it in 0.1-0.2 seconds, or sometimes a bit more.

Human pattern recognition ability is especially strong at picking out edges. Clearly this is a trait with survival value: a leopard's camouflage resembles very closely the pattern of sunlight falling through leaves in a woodland glade. Nonetheless someone used to seeing leopards in the wild will easily pick out a leopard from its background, in less than one second (a city person takes quite a bit longer).

What makes the pattern discrimination possible is the fact that the leopard's spots do not quite match the glade at the edges.

On the other hand, our hard-wired pattern discrimination mechanisms can lead to the eye being fooled by certain patterns. This is the basis of optical illusions such as that shown below:



----all three lines are identical in length!

A final remark about the role of the brain in visual perception: referring to the cross-section of the eye on p. 99, we see that where the retinal neurons (in front of the light receptiors!) join to form the optic nerve, they must pass through the photosensitive layer, thereby creating a "blind spot"----if light falls at that point it cannot produce a visual stimulus. But normally we are completely unaware of this spot. We certainly do not have a blank in our visual fields.

The pattern below verifies the reality of the blind spot:

Block your right eye and look directly at the circle with your left eye. Now move closer to the page: at a certain distance the square will disappear. If you move closer it reappears. Similarly if you block your left eye and concentrate on the square with your right, the circle will disappear as you move toward the page.

The reason you do not notice the blind spot in each eye is that your brain carefully fills in the visual field where the empty spot would have been, using parts of the neighboring image to "dub in", or interpolate the missing information. You may have noticed that when either spot disappeared, it did not leave a hole, but was replaced by the uniform background of the page. That was your visual processing centers at work!