Physica Scripta. Vol. T109, 75–79, 2004

A Short History of Optics

Brian Vohnsen*

Laboratorio de Optica (Dept. Fisica), Campus de Espinardo (Edificio C), Universidad de Murcia, 30071 Murcia, Spain

Received May 1, 2003; accepted June 30, 2003

PACS Ref: 01.65.+g, 42.15.–i, 42.25.-p, 42.50.–p, 42.66.–p.

Abstract

The science of optics stems from the roots of western civilisation and has since then served as a prime source of inspiration for a number of distinguished scientists working towards improving our knowledge about light, colour and the process of vision. In this contribution major steps in the evolution of the intriguing science of optics are described including the various theories that consider light to consist of rays, particles, electromagnetic waves, and lastly a dual nature of light that show both particle and wave characteristics. In route the description includes several of the main players in constructing our present-day understanding of light.

1. Introduction

The Oxford English Dictionary defines Optics as the "science of sight, or of the medium of sight, i.e., light; that branch of physics which deals with the properties and phenomena of light" [1], but this definition poses questions like "what is light?" and "what causes sight?" In this contribution I will highlight what I consider to be the essential cornerstones in the development of the science of optics and the simultaneous search for an answer to the nature of light. I should immediately confess that I will not be able to provide a completely satisfactory answer to this question, as have nobody else to the best of my knowledge, but the exposition should serve to illustrate that although the nature of light may not be fully known, we at least have come to know a great deal more about how light behaves in a variety of situations.

The first notions of light, which can be traced back to several ancient cultures, were of a religious nature. Hence, the sun was considered to be divine as, e.g., Ra, the sun god of the ancient Egypt, from whose eye the light of day was believed to emanate [2]. The bible too presents us with an understanding of light as of divine origin already in the first day of genesis: "God said, 'Let there be light,' and there was light; and God saw that the light was good, and he separated light from darkness. He called the light day, and the dark night." It is noteworthy that the biblical creation of light comes before both the creation of the sun, moon and the stars, and the creation of man. Light is a divine-created entity, as opposed to darkness, and this alone may serve to explain why light is often being associated with good and dark with evil. The Greek philosophers were the first who sought to explain the nature of light and their argumentation was intimately related to the problem of explaining vision. In accord with different held beliefs, they considered the process of seeing as being due to various relations between the eyes of the observer and the object seen. The Pythagorans imagined that something came forth from the eye to the object, the followers of Democritus believed that something solid and extended like a husk, carrying information about the objects shape and colour, reached the eye from the object, and the followers of Empedocles thought that seeing was a combined effect of both something coming forth from the eye to the object and vice versa.

Euclid (~300 B.C.) was the first to state a number of important properties about light that are still commonly used. These include the rectilinear propagation of light rays and the law of reflection (equal angle of incidence and reflection) [3]. When combined with Hero's principle (~100 A.D.), that light follows the shortest path, and the law of refraction, whose origin dates back to Ptolemy (~170 A.D.), one has the essential basis of so-called geometrical optics. The ancient Greeks, however, never succeeded in discovering the mechanisms involved in vision presumably since their thoughts evolved more around geometrical principles, e.g., they considering sight to be restricted to visual cones or pyramids with their apex located at the eye, rather than actual studies of the eye.

Much later, the Arab scientist Alhazen (~1000 A.D.) resolved the puzzle about whether something was being emitted from or to the eye to cause the sensation of vision. His argumentation was based on the observation of pain felt when looking directly at a bright object such as the sun, and the observation of after-images produced after looking for a while at bright objects. As Galen (~200 A.D.) had before him, he described the anatomy of the eye on the basis of dissections although he erroneously concluded that vision was produced at the surface of the crystalline lens. He considered every object point to be a source for straight light rays holding a one-to-one correspondence with the image seen. It is interesting to note that Alhazen's conclusions are mostly based on experimental observations, in contrary to the works of the ancient Greeks, but nevertheless in the west, where the older texts were of stronger influence, the ancient Greeks concept of visual rays that originated in the eye held sway for a long time yet.

Towards the end of the 13th century Italian glassworkers made lenses that were used for spectacles and thus to correct for poor sight caused by presbyopia. This was a major invention since lenses had previously only been used as burning and magnifying glasses and possibly for decorative purposes [4]. This invention, however, did not spur any scientific interest towards explaining how they function so as to improve sight, and such lenses were merely considered magical curiosities of no use for serious studies [5]. Centuries later, Kepler (1604) took up the challenge and studied the transmission of rays through

---

*e-mail: vohnsen@um.es

lenses with geometrical optics, and he applied his studies to the eye. Thus, he was able to describe how spectacle lenses could correct for poor sight and he realised that the external world was seen, not on the crystalline lens as previously thought, but via the inverted image that was formed on the retina. This explanation had earlier, but less precise, been expressed by da Vinci who compared the eye to his camera obscura—a simple instrument consisting of a tiny hole through which light rays pass to draw a picture of the outside world on a distant screen.

Towards the end of the 16th century Dutch spectacle-makers had combined two lenses to facilitate magnified viewing of distant objects, and consequently the telescope had been invented. Galileo (1609) was the first to see the scientific importance of this tool, and he began to produce his own telescopes that had magnifications up to 30 times. He looked at the planets and discovered four of the moons that revolve around Jupiter, and he looked in detail at the phases of the Moon and Venus. Instead of distrusting the images brought to his eye with the aid of this novel instrument he believed in it and found it to support the view of the Copernican (heliocentric) planetary system rather than the commonly believed Ptolemaic (geocentric) system. Kepler saw the importance of these new discoveries and in little time he produced the theoretical explanation for the working principle of lenses and their combined effect in telescopes. He did not have the exact refraction formula at his disposal but only the approximation valid for small angles: that the angle of incidence and the angle of refraction are proportional. We are now at a state that shows a new scientific method based on observations and their comparison with a theory. Although this approach had a hard time to be accepted, it turned out to be what was needed so as to make significant progress when compared to the former methods that were mostly based on assumptions and deductions without the need for experimental confirmations.

In the following sections, I will discuss the outcome of such a scientific approach to explain the nature of light and the mechanisms of sight. Recapitulating, we have at this point that light is an entity that follows rays from every point of an object to the eye of the observer. The object can either be a direct source of light (such as the sun or a fire), or it can be indirect light cast of an object to reveal its form and colour to the observer. The object is seen at its proper spatial location because of the divergence of the rays upon entering the eye pupil, and the image is produced on the retina at the back of the eye from where it is communicated to the brain.

2. Corpuscular or particle theory of light

That light should consist of material particles is an idea that may be argued to have its origin with the atomists of Democritus. This belief gained, at least partially, support by Descartes (~1637), and it reached its zenith in the work of Newton (~1704). Among other models, Descartes had used a mechanical analogy of light by assuming that it consist of tiny particles, and from this hypothesis he had derived the correct expression of the refraction formula. Nevertheless, Snell had experimentally found the same result about 15 years earlier, and it is therefore often known as Snell's law of refraction. With regard to colour, Descartes suggested that it could be a manifest of different angular velocities of rotation of the light particles, and this correctly foreshadowed a kind of periodicity and that colour is a property inherent to the light itself. Previously thinkers had mostly held that light and colour were two completely different entities with colour as a property of the objects that was carried forward to the observer by light. In passing it should be mentioned that Fermat (1679) had formulated a theorem, Fermat's principle of least time (a refinement to Hero's principle of shortest path), from which the refraction formula could be alternatively derived without any mechanical assumptions about the light. There was, however, an important difference between the two derivations since the latter required that the speed of light propagation should be lowest in the denser medium as opposed to the derivation of Descartes where the speed of light ought to be largest in the denser medium. I will return to this point in Section 5.

Newton made a series of wonderful experiments in order to reveal the nature of light. Perhaps the best-known one is where he found that spectral colours can be extracted from white light by making use of the refraction of light when transmitted through a prism. In consequence, he found that white light is constituted by light of various colours and that different colours have distinct refrangibility (i.e., the law of refraction depend on colour). He believed that light was of a material origin consisting of minute particles and that an attraction of these towards the larger body (prism) was the cause of the refraction phenomenon. In such a mechanical picture the different refraction experienced by rays of distinct colour corresponded to light particles of different size. He also studied the partial reflection of light from transparent materials and argued that it could be accounted for by introducing a curious property of light: that it could be in fits of easy transmission or reflection that made it prone to be either transmitted or reflected [6]. Another important study was the so-called Newton's rings observed when having a small but varying air gap between a spherical and a plane surface of transparent material. He observed coloured rings and deduced that their size indicated a 14:9 ratio in the width of the air gap when comparing rings of utmost red and blue [7]. This observation shows that he was in some ways closer to a wave picture of light (see Section 3) than often presumed. Finally, he speculated about another property of light, experimentally discovered by Bartholinus (1669) and later known as polarisation, that light particles could have sides. This, however, he finally left among a number of queries for others to investigate in further detail.

Summarising, the nature of light was now that it consisted of tiny particles, or corpusculars, and that colour was an essential property of the light attached directly to its constituents. No doubt because of the impressive and detailed account of the experiments carried out together with the authority of Newton, established in his previous work on mechanics Principia, Newton's particle view of light became the stronghold for an explanation of the nature of light for about one century. Not everyone agreed, however, on his explanation of light as small particles, as we will see in the following section, or his spectral theory of colours contained in white light. Goethe (~1810) was among those who later attacked the Newtonian view of light and colour, and he held that colours were better considered in relation to how they are perceived. He too made observations with prisms, but without restricting the width of his light beam, and he therefore did not value the colour spectrum that Newton had observed when using a narrow beam of light. His observations revealed more about the response of the visual system when exposed under various conditions, and thus it was an important contribution to physiological optics.

3. Wave theory of light

Already before Newton's work on optics, Grimaldi (1665) had examined transmission of light by a small hole and its passing of tiny obstacles, and he had observed that a small amount of light appeared in regions of expected shadow had it followed straight lines in accord with classical belief. He called this phenomenon diffraction (Newton later called it inflection) and he observed that the light pattern in the vicinity of shadows consisted of complicated coloured bands. He also experimented with prisms and noticed a colouring of the transmitted light beam that, together with other observations, led him to conclude that colour must be a property inherent to light itself [8]. Finally, he speculated whether light propagate in a wave-like fashion but eventually denied the possibility. This, however, was exactly what both Hooke (1665) and later Huygens proposed. An attempted analogy with sound, known to be waves, had previously been used to reject the hypothesis of light waves since it was found not to pass around corners in the same manner as sound. Nevertheless, the diffraction experiments of Grimaldi had already shown that light propagation was not necessarily truly rectilinear.

Huygens (1690) believed in the wave-like nature of light and considered it to propagate as small spherical waves that add up to form a wavefront that afterwards acts as a source for new secondary waves. A continuous repetition of this principle suffices to account for the propagation of light. With this model he was able to explain both the reflection and refraction of light, but a crucial feature was that the speed of light had to be slower in the denser medium. This was in contradiction with the corpuscular theory of light although both had given the correct expression for the refraction formula. Euler (1746) contributed to the theory of light waves and, in analogy with sound, he proposed that each pure spectral colour corresponds to a sinusoidal wave of a unique frequency and wavelength. The downfall of the corpuscular theory of light, however, only gained momentum towards the end of the 18th century. Its most prominent opponent became Young (1803) who, on account of his recent principle of interference, gave an explanation of the phenomenon of Newton's rings in terms of either cancellation or addition of wave amplitudes of light reflected from the two interfaces. He studied various interference phenomena, including the well-known experiment with transmission of light through two closely-spaced small holes, that allowed him to estimate the light wavelength. For utmost red he found $\sim 0.71\,\mu\text{m}$ and for blue $\sim 0.42\,\mu\text{m}$ (the visible range is now often identified with 0.40–$0.70\,\mu\text{m}$) [9]. He also studied vision and found that accommodation was caused by changes in the shape of the eye lens. Moreover, he suggested that the eye has a discrete number of light-sensitive elements with only three kinds of colour responses (red, green and blue–violet) so that other colours are seen via a proper combination of those. These are now known as cones and together with rods they constitute the light sensitive detectors of the retina. Curiously, Newton had previously considered undulatory motion to be a plausible cause for our sensation of colour, and it should be noted that the aforementioned ratio of Newton's rings ($\sim 14/9$) corresponds quite well to the ratio of the upper and lower wavelength limit of the visual spectrum.

A solid basis for the wave theory of light was formulated later by Fresnel (1819) who also did experiments on diffraction of light similar to those by Grimaldi and Newton. By use of a small lens to collect sunlight, and thereby to obtain a more intense source, together with direct observations through an eyepiece, he was able to study the phenomena to an unprecedented level of accuracy. He considered the light to propagate as a sum of Huygens waves that produce both diffraction and interference effects, and he submitted a mathematical description of his model to the French Academy of Sciences that had announced a competition to settle the disagreement concerning a proper theory of light based on either a particle or a wave description [10]. Fresnel won the prize and this proved the end to the former particle description of light. A curious aspect in relation to the competition, however, was that Poisson, who formed part of the evaluation committee, had derived from Fresnel's theory that a bright spot should be produced in the centre of the shadow of an opaque circular disc and this, he believed, would show the wave theory to be erroneous. When the crucial experiment was performed a tiny bright spot could indeed be observed in the centre of the shadow, and there was therefore no other remedy than to surrender to the new theory of Fresnel. It remained to include phenomena of polarisation in the theory and this Fresnel accomplished, together with Arago, by use of a birefringent crystal. They made the two refracted light beams produced by illumination of the crystal (one is the ordinary beam since it obeys the standard refraction formula whereas the other is known as the extraordinary beam) overlap so as to produce interference. No interference pattern, however, could be observed and they concluded that the two waves had to be transverse and orthogonally polarized and not longitudinal waves as previously believed. Finally, a curious phenomenon called conical refraction, which is the propagation of a light cone when a beam is incident along special directions in certain types of birefringent crystals (i.e., along the optical axes in biaxial crystals), was predicted by Hamilton (1832) on the basis of Fresnel's wave theory. Its subsequent experimental verification by Lloyd (1833) strengthened even further the belief in the wave nature of light.

Maxwell (1873) is known for the great achievement of combining all that was hitherto known about the phenomena of electricity and magnetism in a single theoretical framework and a set of four famous equations—the Maxwell equations. From those expressions a wave equation for an electromagnetic field can be derived and he found that the combination of two constants, the vacuum permittivity and the vacuum permeability, led to a remarkable number that coincided (within experimental accuracy) to the speed of light in vacuum [11]. As a consequence, he suggested that light waves must be of an electromagnetic nature. It should be mentioned that Lorenz (1867) too had reached this conclusion on the basis of his work on electrodynamics [12]. Later, Hertz (1888) made experimental work on the emission and receiving of electromagnetic waves (at MHz frequencies and thus well below that of light) and he found that they obey the same reflection and refraction formulas as light and indeed do propagate with the speed of light as previously predicted.

In consequence, light had now been found to belong to a broader class of transverse electromagnetic waves but identified by the wavelength range of the visible spectrum where spectral colours refer to distinct wavelengths. Nevertheless, light is often identified with a broader spectrum beyond what is visible to the human eye so as to include also infrared and ultraviolet radiation discovered as extensions to the visible spectrum already in the beginning of the 19th century.

4. Modern understanding of light

At the turn of the century the wave nature of light was considered an undisputable fact, and the description of matter moved more into focus. Planck (1900), however, studied a subject at the borderline between light and matter as he tried to reconcile classical theory with a description of temperature-dependent electromagnetic radiation from a blackbody (i.e., a perfectly absorbing material). He was soon confronted with a problem that would only give way to the correct solution for the radiation spectrum when he postulated that the internal energy of the blackbody material could only change as an integer number of small quanta. This meant that the electromagnetic field would also change energy in discrete jumps with the size of this energy quantum when radiation was either emitted or absorbed by the blackbody. Einstein (1905) made the bold assumption that light then too consisted of tiny energy quanta, equal in magnitude to those of the blackbody problem. On this assumption he was able to explain the photoelectric effect (i.e., the emission of electrons from metallic plates when exposed to light) that curiously had been discovered by Hertz in his work to confirm the electromagnetic waves predicted by Maxwell (see Section 3). Thus, light interacts with materials as if it consist of particles each carrying a tiny lump of energy. These light quanta were later named photons by Lewis (1926) in analogy with elementary particles. Did this mean a return to the corpuscular theory of light by Newton? No, light is now considered to be of a dual nature with both particle and wave features, where either characteristic can appear dependent on the actual experiment being performed.

During the 19th century another line of research had been the observation of narrow spectral lines of absorption in otherwise continuous spectra and corresponding discrete emission lines of heated chemical substances. Bohr (1913) suggested a model of the atom as a tiny planetary system where the electrons revolve around the heavy atomic nucleus, but with the special feature that each electron is obliged to follow any of a discrete number of orbits. A light quantum of definite frequency can then be emitted from the atom simultaneously with a jump of an electron from one orbit to another allowed one closer towards the nucleus. This classical picture of the atom with a tint of quantum mechanics worked well for the hydrogen atom but could not account for the spectra of more complex atoms. The ideas of Bohr were later put on more solid ground by Heisenberg, Born and others by replacing the orbit concept with probabilities and in this framework light quanta can be emitted from an atomic system each time it relaxes from one state to another of lower energy.

The latest major contribution to an understanding of the nature of light has been the quantum mechanical picture of light contained in quantum electrodynamics developed by Feynman and others. A popular exposition of this theory can be found in Ref. [13]. Light is considered to propagate as a wave function and thereby requiring that all possible propagation paths shall be considered from the emission to the detection of the light. When detecting the light its wave function collapses and the interaction is therefore seen as if light consists of individual particles of light. This dual nature of light is still a highly active area of research being explored, e.g., in single photon experiments.

5. The speed of light

So far we have not ventured into a discussion of the speed of light as I consider this subject of such interest that it warrants a separate discussion. The Greek philosophers considered the speed of light to be very large if not infinite. In the case of visual rays leaving the eye it was apparent that just by opening the eyes one could immediately see distant stars and, in consequence, the speed of light had to be extremely large. Galileo, however, appears to be the first who actually tried to measure the speed of light by sending light signals between two hills. Unfortunately, this did not allow him to conclude anything more than the speed was indeed very large when compared to, e.g., the speed of sound in air. Descartes stated that the speed of light was infinitely large but also used the model of light particles that had different speeds in different materials in order to derive the refraction formula. The first successful measurement of the speed of light was carried out by Römer (1676) at a very large scale when he observed the variation in the period of the moons of Jupiter as observed from the Earth. He concluded that this variation was due to changes in the distance between Jupiter and the Earth at different seasons and that a finite speed of light was the explanation of such a delay. On this basis, he concluded that light takes ~11 minutes to reach the Earth from the Sun (the correct value is ~8 min.). Subsequently, Huygens used the newly estimated Earth to Sun distance to estimate the speed of light as corresponding to ~230,000 km/s. Nonetheless, the determination of this value did not at first influence the choice between a particle or wave theory of light. It was needed to determine whether the speed of light either increased in denser materials (as believed by the supporters of a particle theory) or decreased (as believed by the supporters of a wave theory). The dispute was eventually settled by a measurement carried out by Foucault (1850) where he found that the speed of light propagation was indeed reduced in water as compared to air in agreement with the wave theory and thus in further support of the theory of Fresnel. Finally, the aforementioned studies of Maxwell provided the means to calculate the speed of light on the basis of purely electromagnetic grounds. The speed of light was now well known but not the reference frame. With respect to what did the light move with the determined speed? The aether was believed to constitute the reference frame and this omnipresent substance had been introduced to avoid the mental problem of having truly empty space in vacuum. This substance was essential in the picture held by the scientists that a medium is required to support the light propagation in complete analogy to sound that propagate via vibrations in the material medium. The aether had to be endowed with a number of curious properties that allowed it to comply with a number of requirements. It had to consist of tiny particles that could penetrate the pores of materials and thereby allow the transmission of light. Moreover, if the earth was surrounded by a stationary aether, the aether should be sufficiently dilute so as not to slow down the planetary motion significantly but dense enough to allow the propagation of light. To detect the presence of the aether a viable way would be to determine small changes in the measured speed of light if either measured in the direction of the motion of the Earth through the aether or perpendicular to it. This became the core issue of interferometric measurements carried by Michelson on a number of occasions with ever-improved accuracy towards the end of the 19th century. The result of the experiments, however, was always negative since the expected small change in the speed of light could not be unambiguously determined. Einstein (1905) interpreted this as demonstrating that the aether medium was obsolete, as had been previously speculated by others including Lorenz and Poincaré, and he took the bold step of assuming the speed of light to be independent of the motion of the observer and thereby initiated the development of his theory of relativity. A detailed account of the laborious search for and description of the aether medium can be found in Ref. [14]. Since year 1983 the speed of light has been defined to a constant value of 299,792.458 km/s, as this conveniently eliminates problems towards further improvements in the accuracy of its determination, and in consequence the meter is defined in relation to this as the distance travelled by light in vacuum in a time interval of 1/299792458 s ($\sim$3.3 ns).

Recent studies have addressed, and some even questioned, the constancy of the speed of light. Without venturing into details on those questions that are still areas of active research we may mention two examples where light signals can travel with a speed different from the speed of light. One is the dramatic slowing down of light signals to only a few metres per second [15]. It is important to note that this is a slowing down of light pulses due to a very strong dispersion and therefore not a reduction of the speed of light. Another example is the transmission of signals across tiny barriers via the so-called optical tunnelling effect. Here experiments have indicated the possibility of signal transmissions at speeds higher than the speed of light in vacuum. This is still a controversial issue however with more points to be addressed [16].

6. Conclusion

In this contribution I have described what I consider to be the most essential contributions to the entwined history of the science of optics. Obviously, in any such a presentation, a large amount of material must be omitted and the history is therefore not complete, but nevertheless it should be possible to draw a general picture of the development from it. I have on purpose not ventured into a description of the many applications that the science of optics has produced, as this would suffice for a separate history, but rather focused on studies that have enriched our understanding of the nature of light. In so doing, however, I ought to mention the invention of the laser that was first realised by Maiman (1960) since this has provided researchers with an extremely valuable source of illumination without which many of the latest advances would not have been possible. The laser is also of utmost importance in a number of applications in society of which perhaps the most notable is the revolutionary development of fibre optical information networks during the last decades.

The initially posed questions about what light and sight is have been answered to such a degree that current scientific knowledge essentially permits. There are questions yet unanswered, and the search for a better understanding of vision, light and photons is still far from complete, so certainly the future of optics looks bright.

References

1. "The Oxford English Dictionary," (J. A. H. Murray, H. Bradley, W. A. Craigie, and C. T. Onions, Eds.), (Oxford University Press, Oxford, 1970), vol. 7, p. 163.
2. Zajonc, A., "Catching the Light: The Entwined History of Light and Mind," (Oxford University Press, Oxford, 1995), p. 39.
3. Ronchi, V., "The Nature of Light," (Heinemann, London, 1970), p. 17.
4. Park, D., "The Fire Within the Eye," (Princeton University Press, 1999), pp. 121–124.
5. Ronchi, V., "The Nature of Light," (Heinemann, London, 1970), p. 71.
6. Newton, I., "Opticks," (Dover, New York, 1952), pp. 281–282.
7. Newton, I., "Opticks," (Dover, New York, 1952), p. 210.
8. Ronchi, V., "The Nature of Light," (Heinemann, London, 1970), p. 143.
9. "Great Experiments in Physics," (M. H. Shamos, Ed.) (Dover, New York, 1987), p. 103.
10. "The Wave Theory of Light, Memoirs by Huygens, Young and Fresnel," (H. Crew, Ed.), (American book company, New York, 1900) pp. 79–144.
11. Maxwell, J. C., "A Treatise on Electricity & Magnetism," (Dover, New York, 1954), vol. 20, ch. 20.
12. Keller, O., "Optical Works of L. V. Lorenz," Progr. Opt. **43**, (E. Wolf, Ed.) (North-Holland, Amsterdam, 2002), pp. 267–268.
13. Feynman, R. P., "QED. The Strange Theory of Light and Matter," (Penguin books, London, 1990).
14. Whittaker, E. T., "A History of the Theories of Aether and Electricity," (Humanities Press, New York, 1973), vol. 1: the classical theories.
15. Hau, L. V., Harris, S. E., Dutton, Z. and Behroozi, C. H., Nature **397**, 594 (1999).
16. Chiao R. Y. and Steinberg, A. M., "Tunneling Times and Superluminality," Progr. Opt. **37**, (E. Wolf, Ed.), (North-Holland, Amsterdam, 1997), pp. 345–405.