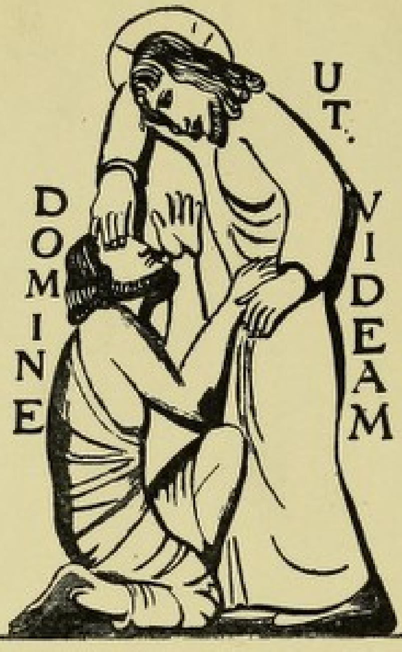

DOMINE
UT. VIDEAM

# THE THEORY OF LIGHT

THE

# THEORY OF LIGHT

BY THE LATE

THOMAS PRESTON, M.A. (DUB.), D.SC. (R.U.I.), F.R.S.

FELLOW OF THE ROYAL UNIVERSITY OF IRELAND; PROFESSOR OF NATURAL PHILOSOPHY, UNIVERSITY COLLEGE, DUBLIN; AND INSPECTOR OF SCIENCE SCHOOLS UNDER THE SCIENCE AND ART DEPARTMENT

*THIRD EDITION*

EDITED BY

CHARLES JASPER JOLY, M.A.

FELLOW OF TRINITY COLLEGE, DUBLIN; ANDREWS' PROFESSOR OF ASTRONOMY IN THE UNIVERSITY OF DUBLIN AND ROYAL ASTRONOMER OF IRELAND

London

MACMILLAN AND CO., LIMITED

NEW YORK: THE MACMILLAN COMPANY

1901

*First Edition printed* 1890. *Second Edition* 1895. *Third Edition* 1901.

# PREFACE TO THIRD EDITION

Soon after the lamented death of Prof. Preston, at the early age of forty, it became necessary to make arrangements for the publication of the third edition of this work. Accordingly, Mrs. Preston placed in my hands a copy of the second edition, annotated by her late husband, together with all his scientific manuscripts. Greatly to my regret, I found that the author had not made any detailed preparations for the new edition. I know that he had contemplated many important changes in the methods of exposition, but of these I can find no trace among his papers, nor can I, from memory, piece them together in any consistent way. He was never in the habit of writing elaborate notes, and, during the last few years of a busy life, his leisure time was fully occupied with his brilliant researches on the electro-magnetic phenomena of radiation in a strong magnetic field. This investigation he prosecuted with untiring activity until it was rudely interrupted by his last illness.

I have therefore been obliged to retain the methods of the last edition, but they are now supplemented by illustrations and examples worked up from the author's manuscripts. These notes are printed without comment, while the additions I have myself made are distinguished by square brackets. The liberties I have taken with the text are few, and, as a rule, so trivial that it would have seemed pedantic to distinguish them in every case.

I am under the greatest obligations to Mrs. Preston for her assistance in the preparation of the new index, and for the great care she has taken in reading the proofs. To Prof. Bergin and Prof. Thrift, F.T.C.D., my best thanks are also due. They have not only read the proofs, but have given me valuable advice in matters of uncertainty. I believe the accuracy of the book will be found to be considerably increased owing to the thorough revision it has undergone.

C. J. JOLY.

THE OBSERVATORY, DUNSINK,
1st *May* 1901.

# PREFACE TO FIRST EDITION

There is perhaps no greater impediment to the advancement of scientific research than the want of an easy channel of communication with all the most recent discoveries. Many of the most valuable of these are hidden in the transactions of learned societies, or scattered in scientific periodicals, published in several languages, and in various parts of the world, so as to be practically inaccessible to many who might otherwise become well qualified to extend the bounds of natural knowledge.

In no branch of Experimental Physics is the English student placed at such a disadvantage as in the Theory of Light, for although we possess some excellent elementary text-books, yet the field covered by them is so limited that they fall far short of the requirements of all who wish to know how far investigation has been carried, or in what directions it remains to be pursued, and of these which are the most urgent and most likely to be attacked with success.

Influenced by these considerations, I have been induced to undertake the present work, with the hope of furnishing the student with an accurate and connected account of the most important optical researches from the earliest times up to the most recent date. I have, however, avoided entering into the more complicated mathematical theories, yet the mathematical theory, in its most elementary form, as well as the experiments on which it is founded, will be found in sufficient detail to enable the student, furnished with the necessary knowledge of higher mathematics, to attack at once with profit the original memoirs and theories recently elaborated by various British and foreign writers.

Thus, although a large part of the book is suited to the reading of junior students, yet I hope it will be found sufficiently full to meet the requirements of those who desire a more special acquaintance with the subject; and to render it more really useful in this respect I have, as far as possible, given reference to original memoirs and other sources whence fuller information may be derived.

The text contains, in addition to the physical theory, a detailed account of the most important experiments and physical measurements, such as the determination of the velocity of light, wave lengths, refractive indices, etc.; and in some of the fundamental experiments, such as those of Newton on the refrangibility of light and coloured rings, I have given extracts from the original accounts, being fully convinced that in power and perspicuity they far surpass any second-hand digest. In this manner I have endeavoured to direct attention to the great importance of Newton's work, and to show that in this department of scientific research also he stands almost without a rival.

Some novelty of treatment will, I hope, be recognised in the extensive application of graphic methods to the solution of problems in diffraction. The calculation of the intensity at the various points of a diffraction pattern by the ordinary methods presents considerable difficulty and labour, but by the method employed in the third section of Chapter IX. almost the whole theory of diffraction is brought within the reach of persons furnished with the most elementary mathematical knowledge, and it might now reasonably form part of the course of very junior students.

An account of the recent, and justly celebrated, experiments of Professor Hertz will be found in the last chapter, together with the mathematical theory of the electric vibrator and the radiation of electromagnetic waves. The importance of these experiments it would be difficult to over-estimate, in so far as they teach us to refer electric and electromagnetic phenomena to the intervention of the same all-pervading medium, which forms the vehicle by which energy passes through space from one body to another, which brings us light and heat from the sun, and to which we now look for a knowledge of the process by which one body is enabled to attract another, as well as for an explanation of the ultimate constitution of matter itself.

In conclusion, I wish to return my best thanks to Professor Wm. Booth (Bengal Education Service), who has been good enough to read through the proofs and make several valuable suggestions. A considerable portion of the proof has also been read by my friends Mr. M. W. J. Fry, F.T.C.D., and Mr. C. J. Joly, so that I trust the work will be found free from any errors or obscurities of a serious nature. To Professor G. F. FitzGerald, F.R.S., F.T.C.D., I am indebted not only for the reading of the proofs, and the most generous assistance and advice, but also for that teaching to which I mainly owe my knowledge of Experimental Physics.

22 Trinity College,
Dublin, *July* 1890.

# PREFACE TO SECOND EDITION

In this edition the text has been revised throughout, and augmented by more than one hundred pages of new matter, in conjunction with which several new diagrams have been introduced. Although these additions are such as to increase the value of the book considerably, yet I must express my regret that, owing to the pressure of other engagements, I have been unable for the present to complete my original design, namely, to bring all parts of the work up to the standard demanded by the present state of science.

Such alterations and additions as I have been able to make are located chiefly in those portions which relate to the rectilinear propagation of light, wave reflection and refraction, and the application of graphic methods to the solution of diffraction problems. More detail has been introduced in some places, especially in the chapter relating to the velocity of light, which now contains an account of Professor Newcomb's valuable experiments.

To my friend Mr. C. J. Joly, F.T.C.D., I am again indebted for his great kindness in assisting me with the proofs.

*March* 1895.

# CONTENTS

## CHAPTER I

### INTRODUCTORY



## CHAPTER II

### PROPAGATION OF WAVES AND COMPOSITION OF VIBRATIONS



## CHAPTER III

### ON THE APPROXIMATE RECTILINEAR PROPAGATION OF LIGHT

## CHAPTER IV

### ON THE REFLECTION OF LIGHT



## CHAPTER V

### ON THE REFRACTION OF LIGHT



## CHAPTER VI

### ON THE DETERMINATION OF REFRACTIVE INDICES

## CHAPTER VII

### INTERFERENCE FRINGES



## CHAPTER VIII

### INTERFERENCE BY ISOTROPIC PLATES

## CHAPTER IX

### DIFFRACTION

## CHAPTER X

### THE POLARISATION OF LIGHT



## CHAPTER XI

### DOUBLE REFRACTION IN UNIAXAL CRYSTALS



## CHAPTER XII

### DOUBLE REFRACTION (FRESNEL'S THEORY)

## CHAPTER XIII

### REFLECTION AND REFRACTION OF POLARISED LIGHT



## CHAPTER XIV

### METALLIC REFLECTION



## CHAPTER XV

### COLOURS OF THIN CRYSTALLINE PLATES

## CHAPTER XVI

### THE STUDY OF POLARISED LIGHT



## CHAPTER XVII

### ROTATORY POLARISATION



### Applications

## CHAPTER XVIII

### ABSORPTION AND DISPERSION



## CHAPTER XIX

### THE VELOCITY OF LIGHT

## CHAPTER XX

### THE RAINBOW



## CHAPTER XXI

### ELECTROMAGNETIC RADIATION

# CHAPTER I

## INTRODUCTORY AND HISTORICAL

### Section I.—Early History

**1. Optics—Definition and Division of the Subject.**—The science of Optics is that branch of natural philosophy which treats of the nature and properties of light and vision. In its domains we meet with a multitude of experiments of exquisite beauty, and investigations which afford ample scope for all the refinements of modern mathematical analysis. It also supplies us with instruments of the highest utility both in the pursuit of scientific inquiry and in the common enjoyments of life. There is accordingly no department of science more deserving of our study, whether we consider the beauty or the multiplicity of its phenomena.

The subject was usually divided by the older writers into *Catoptrics* and *Dioptrics*, which embraced the phenomena arising from *reflection* and *refraction* respectively. These terms have now fallen into disuse, and two branches of the subject have been developed under the titles *Geometrical Optics* and *Physical Optics*. The former is a purely ideal structure built on the assumed truth of the laws of reflection and refraction, with the supposition that light travels through isotropic substances in right lines or *rays*. It is consequently a mathematical development of the two laws by which it assumes the rays to be controlled, and any inquiry into the physical cause or nature of light is outside its province. This inquiry comes within the scope of physical optics, the aim of which is to determine the physical processes concerned in the production and propagation of light, and to account for them by dynamical principles. *Physiological Optics* deals with the phenomena of vision or the sensation produced by light falling upon the retina of the eye.

**2. Ancient Use of Metallic Mirrors and Burning Glasses.**—The ancients appear to have been wholly ignorant of the theory of optics, and to have been exceedingly slow in advancing the construction of optical instruments, almost all the refinements of the subject having been originated and developed within the past three centuries. Yet it cannot be doubted that some of the more striking of the fundamental phenomena were observed and studied in the earliest times of civilisation. If the physical theories of light and vision were subjects too profound for their investigation, they could not fail to acquire some knowledge of the laws of the reflection of light and the formation of images. The attention of the most careless observer must have been attracted with wonder to the image of himself depicted in still water: and the reflected landscape, or the image of a few bushes on the margin of a lake might have afforded to the humblest inquirer an assemblage of observations from which the general laws of reflection could be easily inferred.

Metallic mirrors, and even glass, seem to have been manufactured long before any of the speculations of the ancient philosophers were recorded. They are distinctly mentioned in the Old Testament (Exodus and Job). The invention of burning glasses seems to have speedily followed the art of glass-making. Aristophanes[1] mentions them as early as 424 B.C.

**3. Pythagoras, Empedocles, Plato, and Aristotle.**—The sources from which light is most copiously derived are the sun, stars, and terrestrial bodies undergoing combustion or heated to incandescence. Such bodies we say are *self-luminous* or simply *luminous*, while *non-luminous* bodies are those which are not visible of themselves, but only when illuminated, that is, when in the presence of a luminous body. The former class we say emit light, while the latter do not. We thus distinguish between *luminous* and *illuminated* bodies.

Simple as it may appear to us to regard a luminous body as the source of some influence, which, acting on the eye, excites the sense of sight, much doubt appears to have existed among those who first investigated the subject as to whether objects become visible by means of something emitted by them, or by means of something issuing from the eye of the spectator. According to the opinion of Pythagoras

[1] *Comedy of the Clouds*, Act II. (performed 424 B.C.): "*Strepsiades.* You have seen at the druggist's that fine transparent stone with which fires are kindled? *Socrates.* You mean *glass*, do you not? *Strep.* Just so. *Soc.* Well, what will you do with that? *Strep.* When a summons is sent to me I will take this stone, and, placing myself in the sun, I will, though at a distance, melt all the writing of the summons." (The writing was then traced on wax spread over a solid substance.)

Pliny also mentions globes of glass, which, when held to the sun, produced combustion, and Lactantius (303 A.D.) states that a glass globe, filled with water and held to the sun, could light a fire even in the coldest weather.

(died 540-510 B.C.) and his followers, vision was caused by particles continually projected from the surfaces of objects into the pupil of the eye; while Empedocles (444 B.C.) and the Platonic school maintained that vision was effected by means of something emitted from the eye itself, which, after meeting something else emanating from the object, excited the sense of sight. In the theory of Plato[1] three elements appear to have been necessary to vision. First a visual stream of light or divine fire emitted by the eye itself. These visual rays entered into union with the light of the sun, and the two together, meeting with a third emanation, from the object seen, completed the act of vision.

The doctrine of visual rays, and emission theories in general, was

[1] ". . . and the pure fire which is within us and akin to this they (the gods) made to flow through the eyes in a single, entire, and smooth substance, at the same time compressing the centre of the eye so as to retain all the denser element, and only to allow this to be sifted through pure. When, therefore, the light of day surrounds the stream of vision, then like falls upon like, and there is a union, and one body is formed by natural affinity according to the direction of the eyes, wherever the light that falls from within meets that which comes from an external object. And, everything being affected by likeness, whatever touches and is touched by this stream of vision, their motions are diffused over the whole body, and reach the soul, producing that perception which we call sight. But when the external and kindred fire passes away in night, then the stream of vision is cut off; for going forth to the unlike element it is changed and extinguished, being no longer of one nature with the surrounding atmosphere which is now deprived of fire: the eye no longer sees, and we go to sleep; for when the eyelids are closed, which the gods invented as the preservation of the sight, they keep in the eternal fire.

". . . And now there is no longer any difficulty in understanding the creation of images in mirrors and in all smooth and bright surfaces. The fires from within and from without communicate about the smooth surface and form one image which is variously refracted. All which phenomena arise by reason of the fire or light about the face combining with the fire or ray of light about the smooth and bright surfaces. And when the parts of the light within and the light without meet and touch in a manner contrary to the usual mode of meeting, then the right appears to be left and the left right; but the right again appears right and the left left, when the position of one of the two concurring lights is inverted; and this happens when the smooth surface of the mirror, which is convex, repels the right stream of vision to the left side, and the left to the right" ("The Dialogues of Plato," vol. ii. *Timaeus*, pp. 538, 539, by B. Jowett).

He is speaking of two kinds of mirrors; first the plane, secondly the cylindrical.

Again, p. 561: "There is a fourth class of sensible things, comprehending many varieties, which have now to be distinguished. They are called by the general name of colours, and are a flame which emanates from all bodies, and has particles corresponding to the sense of sight. . . . Of the particles coming from other bodies which fall upon the sight, some are less and some are greater, and some are equal to the parts of the sight itself. Those which are equal are imperceptible, or transparent, as they are called by us, whereas the smaller dilate, the larger contract the sight, having a power akin to that of hot and cold bodies on the flesh, or of astringent bodies on the tongue. . . . Wherefore, we ought to term that white which dilates the visual ray, and the opposite of this black."

combated by Aristotle as early as 350 B.C. He maintained that light is not a material emission from any source, but a mere quality of, or action (ἐνέργεια) of a medium which he called the pellucid (διαφανὲς).[1]

Although the reasoning of Aristotle was very superficial, yet he is entitled to considerable credit for his sagacious, though vague speculations regarding the nature of light and various optical phenomena. He may to some extent be regarded as having in a haphazard manner anticipated the undulatory theory of light, which was established two thousand years afterwards by the labours of Huygens, Young, and Fresnel.

**4. Knowledge of the Ancients.**—The principal phenomena of the rainbow, halos, etc., had not escaped the notice of the ancients, who classed all these appearances under the common denomination of *meteors*. Aristotle[2] attributed these phenomena to the reflection of the sun's rays from drops of rain, and observed that a rainbow may be made by the spray from an oar, and that in this case it will be visible

[1] "There is then, let us begin by saying, something which is pellucid. And by pellucid is meant something which is visible, not visible by itself (to speak without further qualification), but visible by reason of some foreign colour which affects its neutral pellucidity. Of this character are air and water, and also many among solid bodies, water and air being pellucid not in virtue of their qualities as water or air, but because they both contain the same element as constitutes the everlasting Empyrean essence. Light is then the action (ἐνέργεια) of this pellucid *qua pellucid;* and whenever this pellucidity is present only potentially, there darkness also is present. . . . Thus we have shown light to be neither fire, nor body generally, nor even the effluvium or emanation from any body (since even in this case it would be a body of a kind), but only the presence of fire, or something like it, in that which is pellucid ; two bodies being unable to exist at one and the same time within the same space. . . . Darkness in fact is really the removal of such a positive quality from what is pellucid, so that light must necessarily be its presence. Empedocles, therefore, and many others who have followed him, have not described the phenomenon correctly in speaking of light as moving itself, and as coming some time or other without our knowing it into existence between the earth and the surrounding air. . . . And the pellucid itself is also similarly dark, but it is so not when it is pellucid in actuality, but only so potentially ; for it is one and the same element which is at one time darkness and at another time light. . . .

"Colour therefore is not visible without the presence of light ; this indeed we saw was the essential character of colour that it is calculated to set the actually pellucid in movement ; and the full play of this pellucid constitutes light. . . . Vision is the result of some impression made upon the faculty of sense ; an impression which cannot be effected by the colour itself as perceived, and must therefore be due to the medium which intervenes. An intervening substance then of one kind or another there must necessarily be ; and were this intervening space made empty, not only will the object not be seen exactly, but it will not be perceived at all" (*Aristotle's Psychology*, book ii. chap. vii., by Edwin Wallace, M.A., Cambridge University Press Series, 1882).

[2] *Meteor*, lib. iii. cap. ii.

to a person who turns his back to the sun in the same manner as in the case of the natural rainbow.

Notwithstanding the absurdity of the doctrine of *ocular beams*, as it was called, the geometers of the Platonic school were acquainted with two very fundamental points in the science of optics. They taught first, that light, from whatever source it might be emitted, travels in straight lines; and secondly, that when it is reflected at any surface, the angle made with the surface by the incident beam is equal to that made with the surface by the reflected beam.

We thus find them acquainted with the rectilinear propagation of light and with the law of reflection—the two facts in the science which we would naturally expect to have been first discovered.

Epicurus, Lucretius, and the other supporters of the quasi-tentacular theory, although they made few or no experiments, lacked not fertility in hypotheses to account for the common appearances of nature. They all had a confused notion that as we may feel bodies at a distance by means of a rod, so the eye may perceive them by the intervention of light. It is very remarkable that this strange hypothesis held ground for many centuries, and little or no progress was made in the subject till it was established on the authority of Alhazen, an Arabian astronomer, in the eleventh century A.D. that the cause of vision proceeds from the object and not from the eye.

**5. Euclid.**—Shortly after the time of Aristotle the celebrated geometer Euclid (300 B.C.) drew up a treatise on optics, which has been handed down with his geometrical works.[1] However, the work is so imperfect and so inaccurate that some have found it difficult to attribute it to one whose geometry is characterised by such perspicuity and accurate reasoning.

**6. Ptolemy.**—The most celebrated of all the ancient writers on optics was the Egyptian astronomer Ptolemy, who flourished about the middle of the second century. He treated of astronomical refraction and of the increase in the apparent diameters of heavenly bodies when near the horizon. He also drew up tables of the values which he found for the angles of incidence and refraction of a beam of light passing from air into glass and water, but he failed to connect them by any law, like all the subsequent writers of the next fifteen hundred years.

**7. Experiment of Cleomedes.**—Next to a straight stick appearing bent when part of it is immersed obliquely in water, the apparent

[1] (Oxford edition of Euclid's works, 1557.) He endeavoured to refute the Pythagorean, or emission, theory of light, and investigated the apparent place of the image formed by reflection at the surface of a polished mirror.

elevation of a coin, or any other object, placed at the bottom of a cup into which some water is poured, is perhaps one of the oldest experiments depending on refraction. It has been referred to by the oldest optical writers, and especially by Cleomedes, who (50 A.D.)[1] also pointed out that, in the same manner, the air by refraction may render the sun visible when it is somewhat below the horizon.

**8. Previous to Alhazen.**—In addition to what has been mentioned, the ancients had a superficial and fragmentary acquaintance with some of the properties of the rainbow, mirage, and halos, but limited as it was, it far exceeded their knowledge of the other branches of physical science. Their knowledge of the general nature of refraction and of some of its applications was exhibited in the construction and use of burning glasses, which were sold as curiosities in the toy shops, and were probably either glass globes filled with water or balls of glass or rock crystal.

**9. Alhazen.**—After a long interval of inactivity the science of optics was taken up and cultivated with assiduity in Arabia. The first real progress in the mathematical theory was made by Alhazen in the eleventh century. He entered into the anatomy of the eye, and examined the rôle played by each part of it in the production of vision. Besides accounting for twilight he showed that by means of the duration of it the height of the atmosphere might be measured. After describing the eye, he explains how it happens that with two eyes we see only one object, and that we see each object, however small, not by a single ray of light (as was at that time supposed), but by a cone of rays proceeding from the object to the eye.

Alhazen treated largely of *optical deceptions*, both in direct vision and also in vision by reflected and refracted light. In this class of phenomena he ranks what was known as *the horizontal moon*; that is, the increase in the apparent magnitude of the moon, or any other celestial object, when near the horizon. In explanation of this phenomenon he says that we judge of distance by comparing the angle under which we see an object with its supposed distance, so that if the angles under which two objects are seen be nearly equal and if the distance of one be conceived greater than that of the other, the more distant object will be imagined the larger. But the sky near the horizon, he says, is always imagined farther from us than any other part of the concave surface, on account of the range of intervening terrestrial objects by which we judge the distance.[2]

[1] *Cyclical Theory of Meteors* (i.e. *stars*).

[2] This account of the horizontal moon Bacon attributes to Ptolemy. As such it is objected to by B. Porta, *De Refractione*, pp. 24, 128 (Priestley's *History*).

Although the work of Alhazen may have been founded on that of Ptolemy, yet he made such a decided advance in the theory, that for five hundred years or more he was recognised in Europe as the chief authority on the subject.

**10. Vitellio.**—In 1270 Vitello or Vitellio, a native of Poland, drew up a treatise on optics[1] less prolix and more methodical than that of Alhazen on which it was founded.[2] Vitellio attributed the twinkling of the stars to the motion of the air in which the light is refracted, and he remarked that the twinkling is increased when they are viewed through water in gentle motion. He also compiled a table of the angles of incidence and refraction of light at the surface of water and glass, of much greater accuracy than that previously given by Ptolemy.

**11. Roger Bacon.**—Contemporary with Vitellio was our countryman Roger Bacon, a man of extraordinary genius, who wrote on almost every branch of science, yet notwithstanding the pains he took with the subject of optics, he does not appear to have made any advance in the theory which Alhazen had already laid down before him. Great as Bacon undoubtedly was, he was far from being free from the prejudices of his predecessors and contemporaries. Some of the wildest and most absurd of the speculations of the ancients had the sanction of his approbation and authority.

The invention of the magic-lantern has been attributed to Bacon, but it has been much disputed whether he was acquainted with telescopes. Certainly if he was unacquainted with spectacles, telescopes, and microscopes, he anticipated their invention in language more than prophetic.[3]

[1] Published by Risner in 1572, with the work of Alhazen translated from the Arabic, under the title *Thesaurus Opticæ*, Bas. 1572.

[2] Vitellio is said to have at first denied that he had any knowledge of the works of Alhazen, but he afterwards retracted this denial and acknowledged himself a disciple of the Arabian philosopher.

[3] He says: "If the letters of a book, or any minute object be viewed through the lesser segment of a glass sphere or crystal, whose plane base is laid upon them, they will appear far better and larger . . . and therefore this instrument is useful to old men, and to those who have weak eyes; for they may see the smallest letters sufficiently magnified." And again: "Greater things than these may be performed by refracted vision. For it is easy to understand by the canons above mentioned, that the greatest things may appear exceedingly small, and on the contrary; also that the most remote objects may appear just at hand, and on the contrary. . . . And thus from an incredible distance we may read the smallest letters. . . . And thus a boy may appear to be a giant and a man as big as a mountain. . . . So also the sun, moon, and stars may be made to descend hither in appearance . . . and many things of like sort which would astonish unskilful persons" (*Opus Majus*, Jebb's edition, p. 377).

**12. The Introduction of Telescopes.**—Although it would appear from the writings of Bacon, B. Porta, and others, that the properties of some form of telescope were known or suspected, yet the construction and practical applications of the instrument do not appear to have been known and published prior to the year 1608 A.D. If known before this date, the instrument was probably the secret possession of certain individuals who employed it in the demonstration of "natural magic."

Like many other discoveries, it is probable that more than one person had hit upon the idea of the telescope, and had constructed simple forms of the instrument for their own amusement and "curious practices" before any public record of the invention was made. For this reason it is not surprising that the early history of the instrument should have been the subject of a lively debate, and that the invention should have been ascribed to different persons and claimed in different countries. The first person, however, who seems to have independently constructed a telescope, and who at the same time published his discovery, was Hans Lippershey, a spectacle-maker of Middelburg, in the year 1608.

No small share of honour in this matter must be ascribed to Galileo,[1] who, in the following year (1609) (having merely heard that the Belgian spectacle-maker had constructed an instrument by which distant objects were made to appear nearer and larger), at once set to work and independently constructed a telescope for himself. With such skill and ability did he apply himself in this matter that in 1610 he finished an instrument of such excellence that it revealed the satellites of Jupiter, and thus broke the dawn of modern astronomy.

It was Kepler (1571-1630), however, who first reduced the theory of the telescope to its true principles, and laid down the common rules for finding the focal lengths of simple lenses, and the magnifying powers of telescopes.

**13. B. Porta—Camera Obscura.**—At the end of the sixteenth century John Baptista Porta (1545-1615), a Neapolitan philosopher and famous collector of mysteries, published his *Magia Naturalis.* To him the invention of the *camera obscura* is due. He remarked that if light be admitted through a small hole in the shutter of a darkened room, external objects will be clearly depicted on the white wall, in their natural colours; and he added that if a convex lens be placed at the aperture the objects will appear so distinct as to be immediately recognised.

**14. A. de Dominis—The Rainbow.**—In 1611 the true theory of

[1] *Opere,* ii. p. 4.

the primary rainbow[1] was at last arrived at by Antonio de Dominis, archbishop of Spalatro. He showed that one reflection and two refractions in the drops of rain were sufficient to bring the rays which formed the bow to the eye of the spectator. This explanation was either verified or suggested by viewing a glass globe filled with water and exposed to the sun's rays under the same circumstances as the drops of rain. Both the primary and secondary bows were afterwards explained by Descartes[2] on mathematical principles.

**15. Snell and Descartes.**—The next great step was made by Willebrod Snellius.[3] About 1621 he ascertained that when light falls upon the surface of a refracting medium, such as glass or water, the sine of the angle of incidence bears a constant ratio to the sine of the angle of refraction.[4] He died, however, in 1626 without having published his discovery. The law of refraction has been consequently often attributed to Descartes, who first published it in the above form, but not, as Huygens states, without having previously perused the papers of Snell. In his investigations concerning the rainbow Descartes also neglects to mention how far he was indebted to the previous discoveries of Antonio de Dominis.

The speculations of Descartes on the nature of light bear some resemblance to those of Aristotle, and it seems indeed extraordinary that after the lapse of so many centuries, during which the attention of many celebrated philosophers was concentrated on the subject, no real progress had been made in the physical theory of light. Descartes imagined light to be due to a pressure transmitted instantaneously through an infinitely elastic medium filling all space, and colours he attributed to a rotatory motion of the particles of this medium.

**16. Newton and Grimaldi.**—It was still supposed that every refraction of light actually produced colour, instead of merely separating

[1] *De radiis Visus et Lucis*, 1611.

[2] *Spec. Meteorum*, chap. viii.

[3] Professor of Mathematics at Leyden.

[4] Although tables of the angles of incidence and refraction for glass and water had been constructed by Ptolemy and Vitellio, the philosophers who studied them failed to discover the law of refraction which lay hidden in them. Even Kepler (*Paralegomena ad Vitellionem*, 1604) laboured unsuccessfully to derive it from the tables of Vitellio.

Snellius observed that if the refracted ray and the incident ray continued through the point of incidence be intercepted by any line parallel to the normal to the surface at the point of incidence, the length of the intercepted portion of the refracted ray is in a constant ratio to the length of the intercepted portion of the incident ray. The form above, in which it was published by Descartes, is merely the trigonometrical statement of the law arrived at by Snell.

It is remarkable that the law of refraction in its most improved form was arrived at independently by our countryman James Gregory. He is therefore entitled to as much honour in this matter as Descartes.

the colours already existing in ordinary white light, but in 1666 Newton made the important discovery of the actual existence of colours of all kinds in solar light, which he showed to be no other than a compound of the various colours, mixed in certain proportions with each other and capable of being separated by refraction of any kind.

Whilst Newton was making his earliest experiments on refraction Grimaldi's treatise on light[1] appeared, containing an account of many interesting experiments on the effects of *diffraction*, which is the name he gave to a small spreading out of light in every direction upon its admission into a darkened chamber through a small aperture. This spreading out (or *inflection*, as Newton called it) of the light shows that light does bend round corners and deviate from the rectilinear path like sound, but to a very small extent, and it forms the subject of one of the most important branches of physical optics. Grimaldi observed that in some instances the light from one aperture tended to extinguish that from another, yet it cannot be admitted, from the nature of his experiments, that he ever observed any true case of the interference of light.

[1] *Physico-Mathesis de Lumine*, Bonon., 1665.

## SECTION II.—DISCOVERY OF THE VELOCITY OF LIGHT AND DEVELOPMENT OF THE CORPUSCULAR THEORY

**17. Römer—Finite Velocity of Light.**—A new era in the history of optics was registered by the Danish astronomer Olaus Römer, who in 1676 made one of the greatest discoveries in the history of the science—that of the propagation of light in time.[1] Römer[2] was led to this discovery by a series of careful observations on the eclipses of Jupiter's satellites. Each satellite, as it revolves round the planet, disappears behind Jupiter and is hidden from view, or eclipsed, as long as the opaque body of the planet is between us and the satellite. As the periodic time of the satellite is small, its motion is rapid and it disappears almost suddenly, so that the interval of time between two successive eclipses can be estimated with tolerable precision. If this periodic time be known, the dates at which successive eclipses will occur can be tabulated beforehand; but Römer found that the observed times of eclipses did not agree with those calculated in this manner, but that certain inequalities occurred which could be satisfactorily explained only on the supposition that light travels with a finite velocity.

In order to fix our ideas, let us suppose the earth to be stationary and that Jupiter is also fixed, and that the satellite under observation moves round it uniformly with the periodic time T. Under these circumstances the successive eclipses will follow each other regularly at equal intervals of time T. On the other hand, if the earth moves away from Jupiter with a given velocity so that the distance between them increases uniformly, then the interval of time between two consecutive eclipses, as observed from the earth, will be increased from T to $T + \tau$, where $\tau$ is the time required by light to traverse the distance passed over by the earth in the time $T + \tau$. So also, if the earth approaches Jupiter, the time between two eclipses will be diminished in a similar manner. Now, on account of their motions round the sun, the distance between the earth and Jupiter increases during one part of the synodic revolution and diminishes during the remainder. In the former part the periodic time T will have an apparent increase, and in the latter a decrease. This increase and decrease was found by Römer to depend on the rate at which the earth is receding from or approaching

[1] See Mach, *Popular Lectures*, p. 52.
[2] *Hist. et Mém.* x. p. 399. *Ph. Tr.* 1677, xii. p. 893.

to Jupiter, and the inevitable conclusion was that light is propagated with a finite velocity.[1]

The velocity so determined was about 192,000 miles per second.

Considering the enormous rate at which light travels, it is not surprising that Galileo and the Academy del Cimento should have sought in vain to determine it directly. In recent times, however, methods of extreme ingenuity have been devised by Fizeau (1849) and Foucault (1850) for *directly* measuring the velocity of light in air or any other transparent medium. These methods will be fully described in the sequel (chap. xix.), and the results leave no doubt as to the finite speed of light, and fix it at about 186,000 miles, or 300,000,000 metres per second.

**18. Bradley.**—For nearly fifty years after the discovery of Römer no further evidence was adduced to show that the propagation of light was not instantaneous, and the results arrived at by the Danish philosopher were doubted, if not denied, in many quarters. However, in 1728 Bradley[2] discovered what is known as *the aberration of light*, which, like many other great discoveries, was made when the author was in pursuit of another inquiry. Intending to verify some of Dr. Hooke's observations on the parallax of the fixed stars, he observed the star $\gamma$ Draconis at Kew in 1725, and found that it was more southerly than it had appeared before, and on carefully observing it, and other stars, for a long time he found that they all had an apparent motion in space. After much speculation as to the cause of this apparent motion he finally succeeded in solving the difficulty by taking into account the motion of the earth together with the fact that light is propagated with a finite velocity, and the result of his calculations gave a value of this velocity agreeing fairly well with that arrived at previously by Römer. This showed that the direct light of the fixed stars travelled with the same velocity as that reflected from the satellites of Jupiter (see further, chap. xix.).

**19. Energy: its Conservation and Transmission.**—When a material particle is in motion we say it possesses a certain store of energy, which we term kinetic, meaning that this energy is due to the

[1] The time taken by light to travel over the radius of the earth's orbit is about 500 seconds. Römer's estimate was much too high, being eleven minutes. That the inequalities noticed in the eclipses of Jupiter's satellites might arise from the finite speed of light was admitted when Römer propounded his views, but nevertheless it was contested that the observed inequalities might be due to want of uniformity in the motion of the satellite itself. This objection is legitimate, and the astronomical methods alone do not place the question beyond doubt. All uncertainty, however, has been removed by the terrestrial methods devised by Fizeau and Foucault.

[2] *Ph. Tr.* 1728, xxxv. p. 637.

motion of the body. The particle may give up part or all of its energy to another, by collision or otherwise, but when any such transference takes place, the amount of energy gained by one particle is the exact equivalent of that lost by the other. If the energy of motion of any body or system of bodies augments or diminishes, the energy gained or lost must have been abstracted from, or given to, some other system. In this respect energy is like matter. The amount of it in any system can be augmented only at the expense of some other system. That is, energy, like matter, cannot be created or destroyed by any machine or process at the disposal of man. All working engines and animals are mere machines for converting energy from one form to another, or transferring it from one system to another. It is in this sense that we speak of the *conservation of energy*, or the permanence of energy, just as we speak of the conservation or indestructibility of matter. This idea of the impossibility of creating or destroying energy, that is, of its ever disappearing in any system or form without appearing in equal quantity in some other system or form, underlies the whole basis of modern physics, and forms its groundwork, just as the postulated permanence or indestructibility of matter forms the foundation of modern chemistry.[1]

Now there are two methods by which we may communicate energy to a body at a distance—take, for example, the case of a ship at sea. We might fire bullets into it, each bullet carrying a store of energy which it deposits in the ship when it strikes it. By this means we might set the ship in motion. But there is another method by which energy may be communicated to the ship. We may use the water or medium in which the ship floats. We may spend our energy in exciting waves in the water. These waves travelling outwards will break upon the ship and set it in motion, thereby communicating a part of their energy to it.

In the former method each bullet acted the part of a messenger carrying a certain cargo of energy from the person or machine that projected it to the ship. Here we have a transference not only of the energy, but also of the matter which carries it. In the case of the waves, the energy is handed on in succession from one portion of the water (or medium) to the next, while any element of the water merely oscillates about its position of rest. We have thus a flow of energy through the water which affords it a means of transit.

As another example of the transmission of energy through matter, we may consider the case of a mass attached to one end of a rod or

[1] This subject is treated of more fully in the author's *Theory of Heat*, Chap. I. Section VII.

rope. When the other end is held in the hand and twisted, the attached body will rotate so as to free the rod from torsion. Here the energy supplied at one end is transmitted along the rod to the mass at the other. There is a flow of energy along the rod.

Hence if by any means we obtain energy from a source situated at a distance, we are forced to seek for the vehicle by which it is conveyed. Either matter has come to us from the source, carrying the energy associated with it, as in the case of pellets fired from a gun, or else the energy has been successively propagated through some medium existing between us and the source. The probability of the discovery of other methods of the propagation of energy, or even the possibility of conceiving some new method is perhaps a speculation of a purely visionary character, and is certainly beyond our grasp at present. It is well, however, to keep our minds open to the fact that there may be methods of which we have no direct experience, and of which we may, or may not, become cognisant as our knowledge of the material universe increases.

**20. Two Modes of Propagating Energy—Two Theories of Light.**—It having been proved that light travels with a finite velocity, and it being accepted that a luminous body, as such, is the source of some mechanical influence which we call light, and which is necessary to vision, and above all that the phenomena of light and heat are manifestations of energy, the question arises as to how and where this energy exists during the interval between the instant it leaves the luminous body and the instant it reaches the observer. Thus light (or heat) requires about eight minutes to reach us from the sun; how and where is this energy stored during the transit, and by what means is it transmitted from the sun to us? Direct action at a distance is out of the question. We cannot conceive of energy disappearing at the sun and reappearing at the earth after an interval of eight minutes without having been propagated continuously in the interval through the intervening space.

In the present state of knowledge we are acquainted with energy only as associated with matter, so much so indeed that matter has been defined as the vehicle of energy. Consequently two distinct and intelligible methods of representing the propagation and nature of light have been conceived. The first (the emission theory), which was elaborated by Newton, assumes that a luminous body, as such, continually emits small particles, or luminous corpuscles, of extreme minuteness in all directions. These particles are projected from the body and travel through space with the velocity of light, carrying with them their kinetic energy; that is, their energy of motion. This theory accounts at once for such general phenomena as the rectilinear propagation, and reflection, of light; but some of its consequences are quite inconsistent with observed facts. For instance, the doctrine as ordinarily expounded has led to the conclusion that light should travel faster in the denser media, like water and glass, than in the rarer less refracting media, such as air, while experiment proves the reverse. This and other facts which it has failed[1] to explain have been satisfactorily accounted for by the second theory (the wave theory), which supposes light to be due to a periodic disturbance in a medium existing between the luminous body and the eye, and permeating all space. This hypothetical medium is called the *ether*. We are not directly cognisant of it by any of our senses, such as touch, taste, or smell, but nevertheless from the phenomena of light (and electricity) we cannot but be convinced that such a medium exists, and thanks to the labours of scientific men, our knowledge of its properties is rapidly increasing.

According then to the second theory—known as the *Wave Theory*—a luminous body is the source of a disturbance in the ether, which is propagated in waves throughout all space. These waves falling upon the eye excite the sense of vision. They travel with the velocity of light, and carry energy from the body which produces them to that by which they are absorbed.

Before proceeding to the history and development of this theory, which is that now universally accepted, we shall first glance at the emission theory and see how far it will account for the facts.

**21. The Corpuscular or Emission Theory.**—This hypothesis assumes that the sensation produced by light is due to a mechanical action on the retina. It formally states that a luminous body emits minute particles[2] which by their impacts on the retina cause the sensation of vision. Very formidable objections to it are presented at the outset. For corpuscles moving with such an immense velocity as 186,000 miles per second would have an enormous momentum unless their mass be small beyond all conception. Now an exceedingly large number of these particles may be made to act together by concentrating them in the focus of a lens or mirror, and the resultant effect of their impulses might be expected to become visible when subjected to the test of experiment. This apparently easy test of the materiality of light was appealed to by many philosophers. The effects they observed were probably due to extraneous causes, such as draughts

[1] Most of these difficulties may be overcome by introducing suitable hypotheses concerning the nature of the corpuscles.

[2] "Are not rays of light very small bodies emitted from shining substances? For such bodies will pass through uniform mediums in right lines without bending into the shadow, which is the nature of the rays of light" (Newton, *Opticks*, book iii. Qu. 29).

caused by inequalities of temperature, and it is now universally admitted that no effect of the *impulse* of light has ever been perceived.[1] The motion excited in the well-known delicate radiometer of Mr. Crookes is attributed to other causes, and it is to be remembered that in experimenting with this instrument the vane first moves towards the light, indicating an apparent attraction; and it is not until rarefaction is pushed to a certain limit that the motion of the vane is reversed and exhibits an apparent repulsive action of the light.

This is not the only difficulty which besets the theory at the very beginning, for we have seen that the light of the sun is propagated with the same velocity as that of the fixed stars, and that which comes directly to us from these bodies travels at the same rate as that which is reflected by a planet or its satellite. The speed of propagation would therefore appear to be independent of the luminous source, as well as of any subsequent modifications which the light may undergo in the celestial spaces. Now the emissive force required to project material particles with the velocity of light is calculated to be over a million of million times greater than the force of gravity at the earth's surface, and even though this prodigious force were the same for the various independent bodies of the universe, Laplace has shown that if the diameter of a fixed star were 250 times as large as that of our sun, its density being the same, its attraction would be sufficient to destroy the whole momentum of the emitted molecules. M. Arago ingeniously escapes this difficulty by admitting that the molecules may be projected with very different velocities, but that there is only one velocity which is adapted to excite the sense of sight.

22. **Reflection.**—According to the theory of emission each luminous molecule travels in a right line through a homogeneous isotropic medium. Let MN (Fig. 1) be the path of one of these molecules, and let AB be a reflecting surface. As soon as the molecule comes within a certain very small distance from the surface, indicated by the line PQ, it begins to experience the repulsive or reflecting action of the surface. The velocity of the molecule at PQ may be resolved into two components, one parallel to AB and the other perpendicular to it. The former component

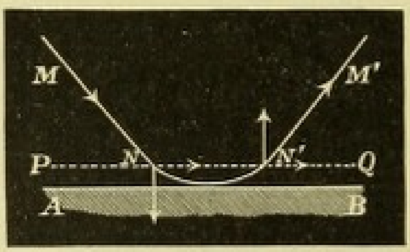


Fig. 1.—Reflection of a Light Molecule.

[1] "The experiments of Mr. Bennet seem to decide this point. A slender straw was suspended horizontally by means of a single fibre of a spider's thread. To one end of this delicately suspended lever was attached a small piece of white paper, and the whole was enclosed in a glass vessel, from which the air was withdrawn by an air-pump. The sun's rays were then concentrated by means of a large lens and allowed to fall on the paper, but without any perceptible effect" (Lloyd's *Wave Theory*).

is unaltered by the action of the surface, while the latter is at first diminished and then reversed, so that the molecule retires from AB at N′ with the same speed as it approached it at N. As soon as the perpendicular component begins to diminish under the reflecting action of the surface the path of the molecule (at N) begins to curve, and when this component is reduced to zero the path of the molecule is parallel to the surface. After this point the repulsive action of the surface will be the same as before, and the route of the molecule will be along a curved path to N′, while at N′ it retires with its velocity perpendicular to the surface reversed, and its velocity parallel to it unaltered. The molecule therefore emerges at N′ free from the influence of the surface in a direction N′M′, making an angle with the surface equal to that made with it by MN. Thus the theory accounts easily for the law of reflection, just as it is deduced for the reflection of a perfectly elastic sphere.

**23. Refraction.**—To deduce the law of refraction from a rare to a denser medium it is assumed that when the molecule comes within a very small limiting distance (PQ) (Fig. 2) of the surface of separation AB, it begins to be attracted towards the surface so that its component velocity perpendicular to the surface gradually increases, till it reaches a limiting distance (P′Q′) on the other side of the surface AB. It then proceeds in the new medium in a right line N′M′, the velocity parallel to the surface remaining the same, while that perpendicular to the surface is increased by an amount which is independent of the angle of incidence, but which varies for different materials. Let the velocity along MN be $v$ and the angle which MN makes with the normal $i$ (the angle of incidence). Then if the velocity along N′M′ be $v'$ and the angle between it and the normal $r$ (called the angle of refraction), we have from the constancy of the velocity parallel to the surface—

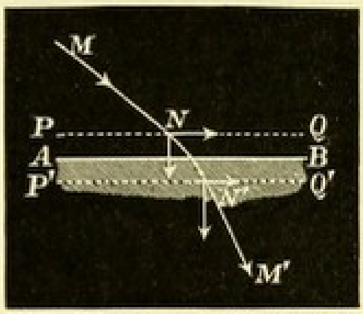



Fig. 2.—Refraction of a Light Molecule.

$$v \sin i = v' \sin r,$$

or

$$\frac{\sin i}{\sin r} = \frac{v'}{v}.$$

The sine of the angle of incidence therefore bears a constant ratio to the sine of the angle of refraction.[1] This is the law of refraction;

[1] If the particle were travelling along M′N′ in the more refracting medium, then in approaching the surface AB it would as before be attracted towards the more refracting medium, so that its component velocity perpendicular to the surface would be diminished by the attraction, and after traversing the curve N′N, it would emerge into the second medium in the direction NM, making an angle with the normal greater than that made by M′N′.

and the formula shows that if $i$ be greater than $r$ then $v'$ is greater than $v$. That is, the velocity of light in denser (more refracting) media is greater than in rare (less refracting) media, for we know by experience that the ray is bent towards the normal (as in Fig. 2) in passing from a rare medium, such as air, to a denser medium like glass or water.

We here reach a crisis in the emission theory, for it has been proved beyond doubt by *direct* experiments on its velocity that light travels faster in a rare medium like air than in a denser (more refracting) medium like water. The emission theory is therefore untenable, and the wave theory, which has not only successfully explained, but even anticipated the results of experiments, has been universally adopted.[1]

**24.** The *prima facie* evidence in favour of the emission theory is very considerable. In the first place it readily accounts for the rectilinear propagation of light, which at first sight looks more like the motion of projectiles than the propagation of undulations which have a tendency to spread out. Then it lends itself at once to the explanation of rays and shadows, while the aberration of light is an immediate deduction. The so-called rectilinear propagation of light was the great difficulty which the early supporters of the wave theory had to face, and the account of it remained in an unsatisfactory state till the time of Young, a hundred years after the time of Huygens who sought for its explanation in certain speculations as to the ultimate constitution of the ether. That no further progress was made until the time of Young has been attributed to the great impulse given to the study of the motion of particles under the action of known forces by the grand discoveries of Newton, which diverted the attention of men of science into that channel rather than to the study of the propagation of undulations.

With regard to the emission theory Sir G. G. Stokes says:[2] "Surely the subject is of more than purely historical interest. It teaches lessons for our future guidance in the pursuit of truth. It shows that we are not to expect to evolve the system of nature out of the depths of our inner consciousness, but to follow the painstaking inductive method of studying the phenomena presented to us, and be content to learn new laws and properties of natural objects. It shows that we are not to be disheartened by some preliminary difficulties from giving a patient hearing to a hypothesis of fair promise, assuming

[1] This difficulty in the theory may be surmounted by a suitable hypothesis concerning the so-called *mass* of the luminous corpuscle.

[2] *Burnett Lectures on Light*, Lecture I., delivered at Aberdeen, 1883.

of course that those difficulties are not of the nature of contradictions between the results of observation or experiment, and conclusions certainly deducible from the hypothesis on trial. It shows that we are not to attach too great importance to great names, but to investigate in an unbiassed manner the facts which lie open to our examination."

**25. Newton's Theory of Fits of Easy Reflection and Easy Transmission.**—The existence of both reflection and refraction at the surface of a transparent substance presents at first sight a great difficulty in the emission theory, for it is not easy to conceive how the same surface may at one time reflect and at another refract an impinging molecule. To meet the difficulty Newton was led from his observations on the coloured rings of thin plates (chap. viii.) to endow the luminous corpuscles with periodic phases or fits, as he terms it, of easy reflection and easy transmission, so that sometimes they are in a condition to be reflected, and sometimes in a condition to be refracted at a transparent surface. To communicate these fits to the luminous corpuscles he imagined all space to be filled with an all-pervading medium or ether. The luminous corpuscles, on striking a reflecting or refracting surface, excite waves in this ether which overtake them at regular intervals, and assist or oppose their motion periodically, so that at any new surface they are refracted or reflected according as the wave assists or opposes the corpuscle. The element of periodicity thus so ingeniously introduced, and which is so fundamentally involved in a wave motion, we should naturally expect to be independent of the angle of incidence. However, to reconcile the theory with his observations on thin plates, Newton found it necessary to suppose the length of a fit to vary as the secant of the angle of incidence, and it does not appear easy to account for such a law.

Boscovich[1] attributed the fits to a polarity of the luminous molecules, which by rotating presented alternately their different sides to the reflecting or refracting surface, and Biot expounded the same theory.[2]

In conclusion, we may state that we believe an ingenious exponent of the emission theory, by suitably framing his fundamental postulates, might fairly meet all the objections that have been raised against it. It will be found, however, on an examination of the whole, that these necessary postulates endow the corpuscles with the periodic characteristics of a wave motion, and when this is introduced the corpuscles themselves may be eliminated, for the wave motion alone sufficiently

[1] Boscovich, *Philosophiæ Naturalis Theoria*, 1758.

[2] Biot, *Traité de Physique*, tom. iv. p. 1.

explains the phenomena. Hence the one remaining argument against the supposition of corpuscles is that they are superfluous, for we believe that no direct test such as has been supposed to be given by the law of refraction in regard to the velocity of light, or by interference phenomena, can decide between the rival hypotheses.

### *Extracts from Newton*

The following passages, quoted direct from Newton's writings, expound his theory in his own words, and show how much more closely than is generally supposed it resembles the undulatory theory now accepted :—

"Were I to assume an hypothesis, it should be this, if propounded more generally, so as not to determine what light is, further than that it is something or other capable of exciting vibrations in the ether ; for thus it will become so general and comprehensive of other hypotheses as to leave little room for new ones to be invented" (Birch, vol. iii. p. 249, December 1675).

*Opticks*, book ii. part iii. prop. xii. : "Every ray of light in its passage through any refracting surface is put into a certain transient constitution or state, which in the progress of the ray returns at equal intervals and disposes the ray at every return to be easily refracted through the next refracting surface, and between the returns to be easily reflected by it.

"This is manifest by the 5th, 9th, 12th, and 15th observations (coloured rings). For by those observations it appears that one and the same sort of rays at equal angles of incidence on any thin transparent plate is alternately reflected and transmitted for many successions accordingly as the thickness of the plate increases in arithmetical progression of the numbers 0, 1, 2, 3, 4, 5, 6, 7, 8, etc., so that if the first reflection (that which makes the first or innermost of the rings of colour there described) be made at thickness 1, the rays shall be transmitted at thicknesses 0, 2, 4, 6, 8, 10, 12, etc., and thereby make the central spot and rings of light which appear by transmission, and be reflected at the thicknesses 1, 3, 5, 7, 9, 11, etc., and thereby make the rings which appear by reflection. And this alternate reflection and transmission, as I gather by the 24th observation (viewing them through a prism), continues far above an hundred vicissitudes, and by the observations in the next part of this book (colours of thick plates) for many thousands, being propagated from one surface of a glass plate to the other, though the thickness of the plate be a quarter of an inch or above ; so that this alternation seems to be propagated from every refracting surface to all distances without end or limitation. This alternate reflection and refraction depends on both the surfaces of every thin plate, because it depends on their distance.

. . . . . . . . . .

"What kind of action or disposition this is ; whether it consists in a circulating or a vibrating motion of the ray, or of the medium, or something else, I do not here inquire. Those that are averse from assenting to any new discoveries but such as they can explain by a hypothesis, may for the present suppose that as stones by falling upon water put the water into an undulating motion, and all bodies by percussion excite vibrations in the air, so the rays of light, by impinging on any refracting or reflecting medium or substance, and by exciting them, agitate the solid parts of the refracting or reflecting body, and by agitating them, cause the body to grow warm or hot ; that the vibrations thus excited are propagated in the refracting or reflecting medium or substance much after the manner that vibrations are propagated in the air for causing sound, and move faster than the rays so as to overtake them ; and that when any ray is in that part of the vibration which conspires with its motion, it easily breaks through a refracting surface, but when it is in the contrary part of the vibration which impedes its motion, it is easily reflected ; and by consequence, that every ray is successively disposed to be easily reflected, or easily transmitted, by every vibration which overtakes it. But whether this hypothesis be true or false I do not here consider."

*Opticks*, fourth edition, 1750, book iii. Qu. 17 : "If a stone be thrown into stagnating water, the waves excited thereby continue some time to arise in the place where the stone fell into the water, and are propagated from thence in concentric circles upon the surface of the water to great distances. And the vibrations or tremors incited in the air by percussion continue a little time to move from the place of percussion in concentric spheres to great distances. And in like manner, when a ray of light falls upon the surface of any pellucid body, and is there refracted or reflected, may not waves of vibrations, or tremors, be thereby excited in the refracting or reflecting medium at the point of incidence and continue to arise there, and to be propagated from thence . . . and are not these vibrations propagated from the point of incidence to great distances? And do they not overtake the rays of light, and by overtaking them successively, do they not put them into the fits of easy reflexion and easy transmission described above? For if the rays endeavour to recede from the densest part of the vibration, they may be alternately accelerated and retarded by the vibrations overtaking them."

And again, Qu. 18 : " . . . Is not the heat of the warm room conveyed through the *vacuum* by the vibrations of a much subtiler medium than air? . . . And is not this medium the same with that medium by which light is refracted and reflected, and by whose vibrations light communicates heat to bodies, and is put into fits of easy reflexion and easy transmission?"

In Qu. 19 he employs this ether (as he calls it) to account for gravitation. "Is not this medium much rarer within the dense bodies of the sun, stars, planets, and comets, than in the empty celestial spaces between them? And in passing from them to great distances, doth it not grow denser and denser perpetually, and thereby cause the gravity of those great bodies towards one another, and of their part towards the bodies ; every body endeavouring to go from the denser parts of the medium towards the rarer? . . ."

Qu. 23 : "Is not vision performed chiefly by the vibrations of this medium, excited in the bottom of the eye by the rays of light, and propagated through the solid, pellucid, and uniform capillamenta of the optick nerves into the place of sensation?" Hearing and animal motion he supposes to be brought about also by the vibrations of the ether.

## Section III.—Introduction and Development of the Wave Theory

**26. Early Speculations.**—The founding of the wave theory of light, like the discovery of the laws of reflection and refraction, has been erroneously attributed to Descartes. In the theory of Descartes vision was supposed to be excited by a pressure transmitted instantaneously through an infinitely elastic medium filling all space, so that it contained nothing analogous to the continuous propagation of waves.[1] The origin of the doctrine might be traced back to the vague speculations of Aristotle, and some germs of it may be found in the writings of Lionardo da Vinci,[2] and in the correspondence of Galileo. More or less obscure ideas were expressed by Grimaldi and Hooke,[3] the latter of whom defined light as "a quick vibratile movement of extreme shortness";[4] but he supposed this movement to be propagated instantaneously in all directions. His theory was consequently little in advance of the instantaneous pressure of Descartes. However, it appears that Hooke was quite prepared to admit that light travelled with a finite velocity (when proved), and that he even anticipated the proof.

The founder of a theory is not, however, the author who makes more or less vague but happy guesses at it, and the credit of discovery is entirely due to him who demonstrates. Otherwise it would be very difficult to fix the date at which the undulatory theory of light was first formulated.

**27. Huygens, Young, Fresnel.**—The true founder of the wave theory is undoubtedly Huygens, who in 1678 first stated it in a definite form, and in 1690 published a satisfactory explanation of reflection and refraction on the supposition that light is due to wave

[1] It is strange that with his ideas as to the nature of heat, which he defines as "an internal agitation of the particles of a body," and though this vibratory motion exists in bodies that are both hot and luminous (*i.e.* incandescent), and is the cause of the "instantaneous pressure" transmitted in all directions, yet there is no statement of a vibration existing in the medium through which the pressure is propagated.

[2] Libri, *Histoire des Mathématiques en Italie*.

[3] *Micrographia* (1665) and *Lecture on Light*. Posthumous works of Hooke, 1705. See p. 76, etc.

[4] *Micrographia*, p. 15.

motion in the ether.[1] He also accounted for double refraction in uniaxial crystals—a phenomenon which had been observed and described by Bartholinus[2] about 1670.

Having failed to account satisfactorily for the rectilinear propagation of light or the theory of shadows, to which the corpuscular theory lent itself so easily, the wave theory, so well initiated by Huygens, fell into disrepute, and remained lifeless for almost a century. It was then revived by Dr. Young's discovery of the celebrated principle of interference.

Although Huygens discovered what is known as the polarisation of light, he was unable to account for it on the wave theory, neither could Young, for these philosophers supposed the wave disturbance in the ether to be longitudinal; that is, in the direction of the ray of light, this being the kind of vibration known to occur in the propagation of sound. And it was not until Fresnel introduced with brilliant success a happy guess of Hooke's (1672), viz. that the light vibrations are transverse—that is, perpendicular to the direction of the ray—that the great difficulties besetting the theory were removed, and the known phenomena not only satisfactorily explained, but others not yet discovered were anticipated. Poggendorff remarks that there is no other instance in the history of modern physics in which the truth was so long kept down by authority.

It was the phenomenon of the polarisation of light that led to the final abandonment of the wave theory by Newton. Having before his mind the longitudinal or sound vibrations, he could not conceive how a ray could have different properties on its different sides. He therefore fell back upon the emission theory, and developed it with a genius more than human.

**28. Interference—Non-Materiality of Light—Experiments of Grimaldi and Young.**—About 150 years before the time of Young, Grimaldi[3] remarked that in certain cases two lights when superposed can partially destroy each other (and Hooke simultaneously laid claim

[1] The only author who can be advanced with any show of reason as an anticipator of Huygens is the Jesuit Pardis. Huygens mentions the manuscript of Pardis, and cites him (*Traité de la Lumière*, p. 18) as "one of those who have commenced to consider the waves of light." The ideas of Pardis were incorporated in the work of another Jesuit, C. P. Ango (*L'Optique divisée en trois livres*, Paris, 1682). It is here explicitly stated that light is due to waves in the ether, just as sound is due to waves in the air.

[2] Erasmus Bartholinus, *Experimenta chrystalli Islandici disdiaclastici*, Copenhagen, 1669; Amsterdam, 1670.

[3] Prop. xxii.: "Lumen aliquando per sui communicationem reddit obscuram superficiem corporis alicunde ac prius illustratam" (*Physico-Mathesis de Lumine, coloribus et iride*, Bologna, 1665).

to the same discovery), but from the manner in which his experiments were conducted he could not have observed any case of true interference. After allowing the sunlight to enter a darkened chamber through two small holes A and B (Fig. 3) pierced very near each other in the shutter, he received the diverging cones of light on a screen. Each depicted a circular spot of light surrounded by a fainter ring. Having placed the screen at such a distance that these rings partly overlapped, he observed that the illumination appeared less in the overlapping

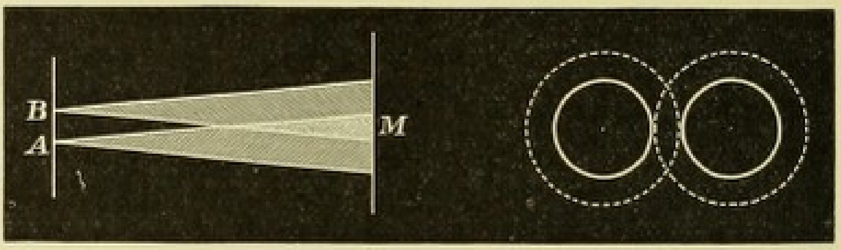



Fig. 3.—Grimaldi's Experiment.

portion than in the remainder of the rings. If one of the pencils was intercepted by an obstacle, this dark portion recovered the brightness of the rest. Thus darkness, he found, may be produced by adding one light to another, and on the other hand, the illumination may be increased by withdrawing a portion of the light. The effect here observed is, however, probably an optical illusion due to contrast, and not a true case of interference.

The object of Grimaldi's inquiry being merely to ascertain whether light was a material or an accident, he prosecuted his research no further, for he considered the experiment fully proved that light was not a material substance.

Young, on the other hand, admitted a very small pencil of light through a narrow slit S in a shutter (Fig. 4). This beam fell upon a screen perforated by two small pin-holes A and B, very near each other. From the apertures A and B he had thus two small pencils of light which he received on a screen M, and he observed

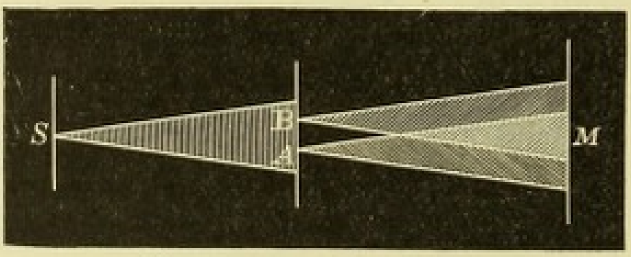



Fig. 4.—Young's Experiment.

that at M, where the pencils overlapped each other, instead of uniform illumination, a series of brilliantly coloured bands appeared (Fig. 5). When he gradually increased the distance between the pin-holes the bands gradually diminished in width till they finally disappeared. They also disappeared when he stopped one of the apertures, or when he removed the slit S and allowed the sunlight to pass through A and B directly, as in Grimaldi's experiment. This showed that the bands must be due to the action of the light from A on that from B, and also that these apertures must be supplied from the same small source S.

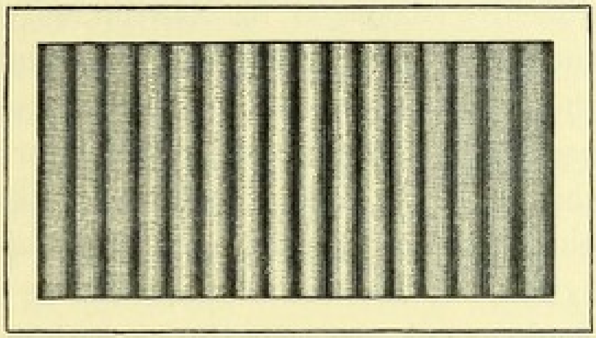


Fig. 5.—Interference Bands.

At any point of a dark band on the screen, the light coming to it from one aperture (A) is apparently destroyed by that coming from the other (B). In this case the two lights are said to interfere destructively. The dark bands are places where the sources A and B produce opposite effects and neutralise each other, whereas in the bright bands the two effects are alike and the illumination is very brilliant. On the whole, however, there is no annihilation of the light. The deficient illumination of the dark bands is accounted for in the excessive brilliancy of the bright bands. The whole quantity of light on the screen is the sum of the quantities which the sources A and B would furnish separately. The dark bands consequently do not point to the *annihilation* of any portion of the light, but merely to a redistribution of it on the screen; and when we speak of destructive interference at any point, it must be remembered that the illumination which is apparently destroyed there exists in equal quantity elsewhere (see further, Art. 44).

No destruction but redistribution.

It is sometimes asserted that the mutual interference and consequent production of fringes by two similar sources of light completely overthrows material or emission theories of light; for, it is said, we cannot conceive of two substances destroying each other. But it should be remembered that here we have no destruction of light, the total quantity remains the same, just as in the case of sand or dust strewn on a vibrating plate or in a sounding tube. If the dust be uniformly distributed on the plate before the vibration starts, it will when the plate is bowed collect along certain lines, leaving the other parts naked. The total quantity of dust however remains unaltered, and the same law holds in Young's interference experiment. This experiment would, consequently, not necessarily overthrow an emission theory, but rather force it to adopt some new hypothesis concerning the corpuscles and their mode of interaction, just as Newton invented his theory of fits to explain the production of coloured rings by thin plates.

Emission theory.

The principle of interference is one of the most fertile in physical science, and many beautiful examples of its power will appear in the sequel. Presently we shall show how it reconciles the apparent rectilinear propagation of light with the wave theory, and answer the difficulty suggested by Newton and those who espoused the emission theory: "If light consists of undulations in an elastic medium, it should diverge in every direction from each new centre of disturbance, and so, like sound, bend round interposed obstacles, and obliterate all shadow." The reply of the wave theory is that light does bend round obstacles (as Newton's own experiments prove), but to a very small extent, on account of the extreme shortness of its wave length, and shadows exist because the several portions of the laterally diverging light destroy each other's effect by interference.

Sound is observed to bend round corners very much more than light, merely because its wave length is vastly greater.

**29. A Medium necessary.**—The radiations which we receive from the sun or from any other luminous body not only affect our sense of sight, which in itself is an evidence of work done, but also in general appreciably heat any body on which they fall. Besides the radiations which affect our sense of sight, and which we term light, a luminous body in general emits others, which we detect by their thermal or chemical action. Heat and work being convertible, we may, by measuring the heating effect of the sun's radiation, calculate the amount of energy transmitted to us per second. We therefore learn to regard the sun or any other luminous body as a source from which energy is emitted in all directions, and the question now arises,—by what means is this energy propagated, and how is it stored while it is travelling to us from the sun or a distant star, for we know that it is not transmitted instantaneously, but travels through the interstellar spaces with a definite velocity, viz. the velocity of light.

Now, with our present experience, it does not seem possible to conceive of more than two modes by which any body as a source of a mechanical influence, travelling with a finite velocity, can ultimately affect and communicate energy to another body situated at a distance. A mechanical influence implies the intervention of a substance of some kind, and this substance may be either projected forth from the influencing to the influenced body, like bullets from a gun, each particle travelling with a certain velocity and carrying a definite amount of energy with it; or it may exist as a continuous medium filling the space between the two bodies, and, being disturbed by one, the disturbance may be propagated from portion to portion of the medium (each part being agitated by its predecessor, and in turn yielding to the succeeding element the energy it received) till it finally reaches the second body. The waves excited by casting a stone into water or by a sounding bell illustrate the latter method of propagating energy, and this forms the basis of the wave theory of light.

Young's discovery of the so-called destructive interference of two lights suggests that light in itself is not a substance emitted by the luminous body. In addition the emission theory has failed (or requires interminable patching) to account for the observed facts. Scientists have consequently been compelled to have recourse to the wave theory, and to assume that all space is filled with some medium or substance, if we may so call it, differing in its properties from visible material bodies, and that a luminous body, as such, is the source of a periodic disturbance of some kind which is propagated in all directions by means of this medium, or, as it is called, *the ether*.[1]

The balance of experimental evidence is in favour of the theory that all optical and radiation effects are due to rapid periodic changes of some properties of the ether. Electric, magnetic, and electro-magnetic effects also appear to be due to the intervention of the same medium.

We know that sound travels through air, water, glass, and other material substances with a definite velocity in each, and experiment proves that the propagation of sound in these substances is due to an undulatory disturbance or vibration, excited in them by the sounding body. Sound is not propagated in a vacuum. Its phenomena are consequences of the vibratory motion of the parts of the material substances through which it travels. When a musical note is sounded energy is transmitted to the air by the sounding body and the air is thrown into vibration. This energy travels through the air as a wave motion, and part of it is spent in exciting the tympanum of the hearer. Light, on the other hand, is propagated with the greatest facility through the best vacuum we can procure. It traverses the interstellar spaces where we cannot suppose any material substances to exist, except perhaps the most excessively attenuated atmospheres or sporadic groups of meteorites. It is propagated with more or less facility through transparent substances, but in all cases with a velocity enormously greater than that of sound. Hence although the presence of material substances modifies the propagation of light to some extent (as by refraction or diminution of velocity, etc.), yet they are by no means necessary to its conveyance from one part of space to another.

The existence of some medium filling all space as far as the farthest star becomes therefore a necessity in the rational explanation of the phenomena of heat and light, and before proceeding to the development of the wave theory of light, it will be well to consider

[1] See further, *Theory of Heat*, pp. 51, 56.

some of the fundamental properties of this hypothetical medium—the ether.

**30. The Ether.**—The assumption of the existence of a medium filling all space does not seem to have presented any serious difficulty to the reception of the wave theory. A far more formidable difficulty with which the early supporters of the theory had to contend was presented by the existence of rays and shadows. They could not explain by the wave theory the apparent rectilinear propagation of light to which the emission theory lent itself so easily.

Several ethers have been postulated by different philosophers for different purposes.[1] Newton supposed that a medium existed in which his luminous corpuscles travelled, and in which they were capable in certain cases of exciting undulations. He also attempted to account for gravitation by the differences of pressure in an ether, but he published little of his theory, "because he was not able from experiment and observation to give a satisfactory account of the medium and the manner of its operation in producing the chief phenomena of nature."

The only ether which has survived is that conceived by Huygens to account for the propagation of light. The evidence in favour of it has accumulated with each discovery of science, and the properties of it as deduced from the phenomena of radiant light and heat are also those required to explain the phenomena of electricity and magnetism. It may be that this same medium forms the vehicle by which gravitation is maintained between material substances, and in some manner as yet unknown to us forms the link of connection by which the sun is enabled to attract the earth and planets and keep them in their orbits.[2] The present tendency indeed of physical science is to regard all the

[1] To Descartes the bare existence of bodies apparently at a distance was a proof of the existence of a continuous medium between them, for he regarded extension as the sole essential property of matter, and matter a necessary condition of extension. "Ethers were invented for the planets to swim in, to constitute electric atmospheres and magnetic effluvia, to convey sensations from one part of our body to another, till all space was filled several times over with ethers" (J. C. Maxwell).

[2] "It is true that, notwithstanding the labours of various scientific men, we are not in a condition to give an explanation of gravitation, but our inability to explain it by no means proves that it is a primary property of matter, incapable of explanation, or forbids us to suppose that it may in some way be brought about by the intervention of that same substance which we find it necessary to assume for the explanation of the phenomena of light on the theory of undulations. . . . Assuming for the moment, as a thing at the present day resting on evidence quite overwhelming, that light consists of undulations, we cannot fail to be impressed by the multiplicity of purposes all bearing so intimately on our wellbeing, which, it seems probable, or not unlikely, are fulfilled by one and the same substance, endowed with properties which we are only gradually learning" (Stokes's *Burnett Lectures*).

phenomena of nature, and even matter itself, as manifestations of energy stored in the ether. When we electrify a body a certain amount of energy is expended, and this is ordinarily regarded as the energy of the electric charge, and may be recovered at any time by discharging the body. But where is the energy stored? We say it is stored in the ether. So again it may be that the energy spent in raising a mass from the earth's surface is stored in the ether.[1] Hence what we call potential energy may be energy stored in the ether, and if it exists there as motion of the ether, then we may regard all energy as kinetic.

To account for the propagation of undulations with a finite velocity and carrying energy, the ether has been endowed with the two radical properties of elasticity and density, or rather something corresponding to elasticity and density.[2] When sound is propagated through material substances, rarefactions and condensations are produced, and to the forces of restitution called into play the propagation of the sound is due,[3] while the velocity of the propagation depends on the elasticity and density of the substance. There is, however, a series of phenomena in light[4] which have no counterpart in the theory of sound, and which lead to the conclusion that the so-called elasticity of the ether is very different from that of the air. They suggest that the vibration of the luminiferous ether must be transverse to the direction of propagation of the light. Air and fluids cannot transmit transverse vibrations, for they offer no resistance to distortion, and this is the property on which the propagation of transverse vibrations depends. When sound is travelling through air the vibrations of the air are longitudinal, that is in the direction in which the sound is travelling. Solids, on the other hand, are capable of transmitting both kinds of vibrations, but with a velocity enormously less than that of light. The elasticity of the ether has consequently been assumed to be somewhat of the nature of that of an elastic solid, but the propagation of light by it on this hypothesis is encumbered by several difficulties. The first is the possibility of longitudinal vibrations or undulations normal to the wave front, as in the case of sound propagation. That no optical phenomena arise from these has been accounted for by supposing the ether incompressible, so that the velocity of propagation of

Transverse vibrations.

Elastic solid theory.

[1] See further, *Theory of Heat*, p. 76.

[2] On this point see *Theory of Heat*, p. 52.

[3] In the case of a vibrating elastic solid the energy is half in the form of kinetic energy, due to the vibratory motion of the parts of the body, the other half being potential; that is, stored up in the distortion of its parts (see Thomson and Tait's *Natural Philosophy*, or Love's *Theory of Elasticity*).

[4] The polarisation of light.

the longitudinal wave is infinite. Again, the phenomena of polarisation and double refraction have led to incongruities and artificial assumptions.

Since the vibrations of transparent bodies travel much too slowly to allow us for a moment to suppose that the propagation of light might be due to them, we are forced to conclude that the ether which conveys the light is distinct from these transparent media, and interpenetrates them all freely (and probably opaque bodies too). It may be difficult at first to admit that a solid body like glass could possibly have ether freely pervading it; but we must remember that the ether is a medium about which our senses give us no direct information. We cannot see, taste, or smell it. It is only by the intellect that we become convinced of its existence,—by studying the phenomena of nature, and finding how they may be explained by it. A magnet attracts a piece of iron even though a plate of glass be interposed between them, and yet the magnetic influence is one which does not directly affect our senses, but we must conclude that whatever medium enables the magnet to put the iron in motion, and communicate kinetic energy to it, permeates the glass as well as the air and interstellar space. This medium also propagates light and heat through many solid substances; it therefore not only interpenetrates them, but is capable of vibrating within them.[1]

Matter and ether.

Now although we suppose the ether to freely permeate all bodies, yet we must suppose that its vibrations are controlled to some extent by the matter of these bodies, for we can prove that light travels with different velocities in different transparent substances (a fact attested by the refraction or bending of the ray in passing from one transparent substance to another), while in opaque bodies it is not propagated at all. It is therefore natural to inquire if the free motion of the ether is influenced by the presence of masses of matter or by matter molecules. Do bodies in motion in this ocean of ether carry with them the ether they already contain, or do they allow the ether to pass through them freely, as water would pass through a net, or, as Young imagined, like the air through a grove of trees? Or does the ether offer any resistance to the motion of the earth and planets through it, and do these bodies, by their motion, produce streams and eddies in it?

The whole question of the state of the ether near the earth, and of its connection with ordinary matter, is still far from being settled by experiment (see chap. xix.).

[1] This free interpenetration follows as a natural consequence of the vortex-atom theory of matter (*Theory of Heat*, p. 77, etc.).

Whatever difficulties we may have in forming a consistent idea of the constitution of the ether, it cannot be doubted that the interstellar spaces are filled with an all-pervading medium in which the ultimate particles of matter and of our own bodies are continually bathed, and yet of which our senses afford us no direct cognisance. And whatsoever other functions appertain to it, one of the chiefest is the conveyance from the sun to his system of that energy by which all its physical life is sustained.

**31. The Vibration.**—Although all the phenomena of interference afford us the strongest evidence that light is propagated as a vibration of the ether—that is, a rapid periodic change of its state or of some of its properties—yet the medium being hypothetical, we are almost wholly ignorant as to what it is that vibrates or how it vibrates. The direction to which the periodic change of state is related we term the direction of the vibration, meaning by the vibration that periodic change, whatever it may be. If it be a periodic displacement, then the direction of the vibration is the direction of the displacement, but this does not necessarily coincide with the direction of propagation of the disturbance. In the case of sound we know well the nature of the disturbance and the properties of the vehicle. In the case of light we shall see that at least a component of the vibration is in the wave front, or transverse to the direction of propagation.

**32. Special Forms of the Wave Theory.**—Many of the phenomena of light, such as the colours of thin plates, can be equally explained by any form of wave theory, and cannot be substantially modified by a more exact knowledge of the constitution of the ether. However, in other parts of the subject, such as the theory of crystalline refraction and reflection, we are compelled to make some hypotheses, and if on developing the special theory built on these hypotheses we find a discordance with any observed phenomena, this disagreement only disproves the special form of theory admitted, but the wave theory in its general sense remains intact.

The necessity of transverse vibrations, and the incapability of fluids to propagate them, led to the development of the most celebrated special form of wave theory—that which regards the ether as an elastic solid. Here, however, we need not regard the ether as possessing all the properties of an elastic solid. All it requires is torsional rigidity, that is, resistance to change of shape, or some property analogous to the torsional rigidity of elastic solids, to enable it to transmit the transverse vibrations of light. A substance like a jelly could transmit either transverse or longitudinal vibrations, and the velocity of the latter might be very great compared with the former. To avoid the difficulty introduced by the possible longitudinal waves, Green and the promoters of the elastic solid theory supposed the ether to be incompressible, so that the longitudinal disturbance travelled with an infinite velocity. Sir G. G. Stokes has, however, remarked that although the ether may act as a perfect fluid for any finite displacements, yet the displacement occurring in the propagation of light may be so small and so rapid that for it the medium behaves as an elastic solid. Cumbered with difficulties which only increase as it is required to meet the demands of accumulating information, the elastic solid theory, although in many instances it comes near the truth, can in optics be regarded only as a first speculation, but nevertheless it must always retain a high historic interest.

The theory which promises most favourably at present is that which regards the ether as a turbulent fluid, and light as an electro-magnetic phenomenon arising from very rapidly alternating electric polarisations or "displacements," as Maxwell termed them. When a body is electrified energy has been spent in producing the electrification, and this energy is stored in the ether around the body. To indicate this we may say that the ether is polarised, meaning thereby that its elements have suffered some directed transformation in properties, or change of state, by the storing of the energy. Electric phenomena are manifestations of this energy in transformation, and when the body is discharged the ether is released from the energy and the consequent polarisation. Similar remarks apply to an electric current. Now if a body be rapidly charged and discharged, or a current be passed rapidly in opposite directions, the ether around will be as rapidly thrown alternately into opposite states of polarisation, and when this becomes very rapidly periodic we have the vibration of the ether spoken of in the wave theory of light. When the rapidity of these vibrations lies within certain limits (red and violet) they affect the eye with the sense of sight; below the red we detect them by their thermal, and above the violet by their chemical action[1] (see further, chap. xxi.).

Apart from its probable truth, the electro-magnetic theory of light shows us how careful we must be to avoid limiting our ideas as to the nature of the luminous vibration.

[1] A quasi-rigidity might be conferred on the ether by other motions going on in it. Thus by filling it with vortices it might become capable of propagating transverse waves and standing electric stress.

# CHAPTER II

## THE PROPAGATION OF WAVES AND THE COMPOSITION OF VIBRATIONS

**33. Wave Motion.**—The essential characteristic of wave motion is that a periodic disturbance is handed on successively from one portion of a medium to another. Examples of it are constantly presented to the notice of every one, as, for instance, when a stone is cast into still water or when a sounding bell throws the air into vibration.

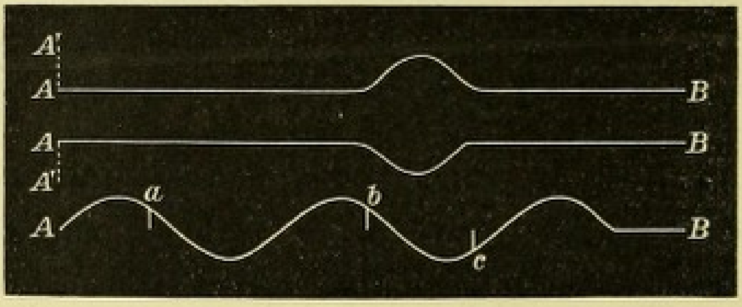
That which is propagated from one part of the medium to the other is energy, not matter, for while any element of the medium merely oscillates about its position of rest, there is a continuous handing on, or flow, of energy from one part to another. In the case of a projectile or a current, on the other hand, matter flows from one place to another, and carries with it its associated energy, so that we have a flow both of energy and matter.

If the medium be homogeneous and isotropic, a disturbance is propagated with the same velocity in all directions; but if the medium be not homogeneous the speed may vary from point to point, and if it be homogeneous but not isotropic the speed may depend on the direction of propagation.

In general the velocity of propagation depends on the nature of the medium and the length of the wave, but so far as light is concerned the velocity in interstellar spaces seems to be the same for all wave lengths.

**34. Transverse Wave.**—As an elementary introduction let us consider the nature of a wave of transverse vibration and its mode of propagation.

Take a flexible cord AB, one end B being fixed and the other A free. (A thick piece of india-rubber tubing 3 or 4 yards long answers very well.) If the free end A be quickly moved from A to A′ (Fig. 6, *a*) and back again, the displacement communicated to A will run along AB as a wave. Fig. 6 (*a*) represents this wave travelling along AB. If A had been displaced in the opposite direction the wave would travel as in ($\beta$), whereas if A were displaced backwards and forwards continuously, a continuous series of waves would be propagated along it as in ($\gamma$). The experiment ($\alpha$) may be easily reproduced by giving a sudden jerk to one end of a rope lying on the ground. In all these cases we speak only of what happens as the wave




($\alpha$) Solitary wave.
($\beta$) " "
($\gamma$) Consecutive waves.

Fig. 6.—Transverse Waves.

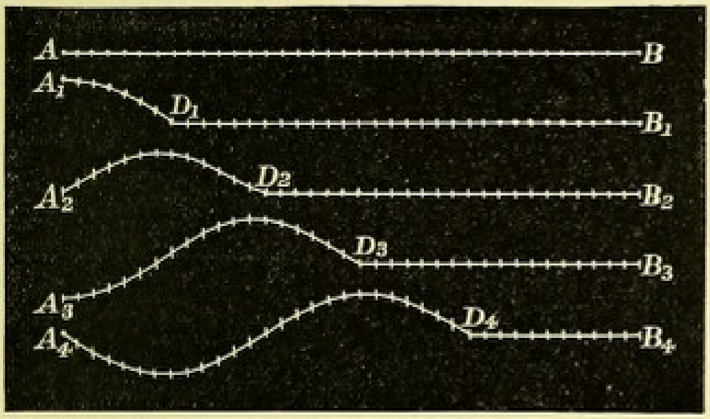
travels from A to B, that is, we suppose B very far away, for at present we are not concerned with what happens after the wave reaches it.

Here it is evident that each element of the cord merely oscillates backwards and forwards, like A, through its position of rest, while the disturbance is propagated from A towards B. There is a flow of energy along the cord.

*Definition.*—Two particles such as *a* and *b* [Fig. 6 ($\gamma$)] are said to be in the same phase of motion when their displacements and direction of motion are the same, and two particles in the same phase are separated by a complete wave length, or by any whole number of wave lengths. Two particles such as *a* and *c*, whose displacements and directions of motion are opposite, are separated by half a wave length, or any odd number of half waves, and are said to be in opposite phases of motion.

Phase.

The process which takes place in the cord may be illustrated by supposing AB a row of particles connected by elastic bands. When A is displaced it drags the adjacent particle after it, which in turn acts on the third, and so on, the disturbance being handed from one particle to the next. Thus by the time A has reached its greatest distance $A_1$ (Fig. 7, $\beta$) from its position of rest, the disturbance will have travelled along the line to some point $D_1$, and the particle at $D_1$ will be on the point of beginning to move as A did. In doing so it will pull its successor after it, and the disturbance will travel on in this way to the end of the line. But when A begins to return to its initial position it drags the adjacent particle after it. This in turn reacts on the next, and so on, so that by the time A reaches its initial position the original disturbance (which was being propagated all the time along the line) has reached a distance $D_2$ equal to twice the distance of $D_1$ from the end A of the line. The row of particles being now represented by Fig. 7 ($\gamma$), let $A_2$ pass through its position of rest to an equal distance $A_3$ on the other side. The configuration of the line is now represented by ($\delta$), in which the original disturbance is just about to displace $D_3$ at




($\alpha$) Connected particles.
($\beta$) Quarter wave.
($\gamma$) Half wave.
($\delta$) Three-quarter wave.
($\epsilon$) Complete wave.

Fig. 7.—Transverse Waves.

a distance equal to $3AD_1$ from A. So again when A returns to its position of rest the line will be represented by Fig. 7 ($\epsilon$), the disturbance having now reached $D_4$ while A has made a complete excursion. $AD_4$ represents a complete wave. It is the distance which the disturbance travels along the line while the particle A makes a complete vibration. Now $D_4$ is just about to repeat the oscillation performed by A, so that if A vibrates continuously $D_4$ will be moving in exactly the same manner and have the same displacement and direction of motion as A at every instant. $A_4$ and $D_4$ are in the same phase, and the length $AD_4$ is a wave length. Thus at distances of a wave length, or any multiple of a wave length apart, the motions of two particles are exactly similar, while at distances of half a wave length apart, or any odd number of half waves, the displacements and velocities are equal but in opposite directions. (This remark is of great importance in the theory of interference.)

Wave length.

**35. Plane Wave.**—Instead of a single cord let us imagine an infinite number of parallel and similar cords with their ends attached to the same plane. For the sake of clearness let us suppose the plane to be horizontal and the cords to hang vertically from it. Then if the plane be moved parallel to itself, so that each point of it describes a short horizontal line, the ends of the cords, being attached to the plane, will be displaced horizontally and a transverse disturbance will run along each. At the end of any time this disturbance will have travelled to the same distance along each cord, and therefore the locus of those points which are just about to be disturbed, or the *wave front*, will be a horizontal plane. If the plane be caused to oscillate regularly backwards and forwards a series of similar waves will run along each cord.

We might now suppose the cords so numerous that they touch each other, and we are thus furnished with an idea of a continuous medium, disturbed by a series of plane waves propagated through it.

The end of each cord has been supposed for simplicity to oscillate along a straight line, but we might equally have supposed it to describe any curve—thus in the case of waves in water the various particles generally describe circles or ellipses in vertical planes, parallel to the direction of propagation.

**36. Spherical Wave.**—Let us now take the case of an infinite number of exactly similar elastic cords attached to the same point, and diverging in all directions from it. Further, let a disturbing force at the centre, acting in the same manner upon each, cause them all to undulate alike. The cords being similar in all respects, it is obvious that the waves propagated along them will be alike and travel with the same velocity. Thus all points at the same distance from the centre will be in the same state of disturbance at the same instant—that is, the locus of points in the same phase of vibration is a sphere and the wave fronts or wave surfaces are spherical. If we imagine the cords to fill all the space around the centre we have the case of a continuous medium disturbed by a system of spherical waves diverging from a point.

*Definition.*—The continuous locus of those points which are in the same phase of vibration is called a *wave front*. The word continuous is inserted because in oscillatory motion such as we are considering a system of successive waves are in similar phases and a succession of similar wave fronts coexist, each a wave length in advance of its predecessor. In this sense any *surface of equal phase* is a wave front.

Wave front.

The wave front might be defined more closely as the continuous locus of those points which are just on the point of being disturbed, and in this sense the term is frequently used. This wave front then marks the limits to which the disturbance has just reached at the instant considered.

## *Examples*

1.—Prove that the equation

$$y = a \sin \frac{2\pi}{\lambda}(vt - x)$$

represents a wave disturbance in which $v$ is the velocity of propagation, $\lambda$ the wave length, $y$ the displacement of a particle from its position of rest at the time $t$, and $x$ the distance from the origin of the same particle.

For brevity denote the angle $(vt-x)2\pi/\lambda$ by $\theta$; then if we change $x$ to $x\pm\lambda$ we change $\theta$ to $\theta\mp 2\pi$, and the sine of the angle remains the same, therefore the value of $y$—that is, the displacement of the particle under consideration—is exactly the same as that of the particle at a distance $\lambda$ from it, or any number of times $\lambda$ from it. Hence $\lambda$ is the wave length.

Again let $t$ become $t\pm\lambda/v$, then $\theta$ becomes $\theta\pm 2\pi$, and therefore the value of $y$ remains the same as before. The value of $y$ for the same particle is therefore periodic, and the *periodic time* T is $\lambda/v$—that is, $\lambda=v\mathrm{T}$. But $\lambda$ the wave length is the distance through which the disturbance is propagated in the periodic time T, consequently the quantity $v$ is the velocity of propagation.

The quantity $a$ is termed the *amplitude* of the vibration. It is evidently the greatest displacement of any particle from its position of rest. The amplitude depends upon the nature of the medium and the power of the disturbing centre, and its magnitude must here be left an open question. The periodic time T can be easily determined in the case of sound, but in the case of light the vibration is so exceedingly rapid that it can only be determined indirectly from the equation $\lambda=v\mathrm{T}$, by means of a previous knowledge of $v$ and $\lambda$.

If we consider a row of particles AB, connected as in Fig. 7 and initially in a right line, their simultaneous positions when disturbed by an undulatory motion will be represented by Fig. 6 ($\gamma$). This curve is determined by the above equation when we suppose $t$ to remain constant and $x$ and $y$ to be the abscissa and ordinate of the particle. But if we confine our attention to a single particle—that is, suppose $x$ constant while $t$ and $y$ vary—then the same curve will determine the displacement $y$ of the particle at any time $t$, the time $t$ being now the abscissa of the curve.

Thus the curve which represents the simultaneous displacements of all the particles also exposes the history of the displacement of a single particle.

*Simple Harmonic Motion.*—If a particle moves subject to the above equation, it is said to execute a simple harmonic vibration. The motion here described is that executed by a particle constrained to move on a right line under the action of a force varying directly as the distance. Such would be the motion of a particle placed in a smooth straight tube passing in any direction through the earth.

2.—If a particle moves round the circumference of a circle with uniform velocity, the foot of the perpendicular from it on any diameter moves backwards and forwards along that diameter with a simple harmonic motion.

Let the angular velocity of the particle be $\omega$, and let the time be reckoned from

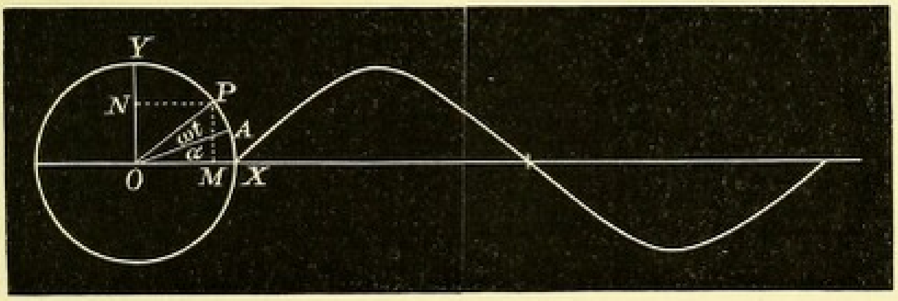



Fig. 8.

the instant the particle leaves the extremity A (Fig. 8) of some diameter OA making an angle $\alpha$ with OX. Then if P be the position of the particle at any time $t$ the angle $\mathrm{AOP}=\omega t$ and $\mathrm{XOP}=\omega t+\alpha$. Hence, if OY be at right angles to OX, the distance $y$ from O of the foot N of the perpendicular from P on the diameter OY is

$$y=a\sin(\omega t+\alpha),$$

where $a$ is the radius of the circle, and is the amplitude of the vibration of the foot of the perpendicular. If T be the time of a complete revolution we have $\omega T = 2\pi$, so that the equation may be written

$$y = a \sin\left(\frac{2\pi t}{T} + \alpha\right).$$

Or substituting from the equation $\lambda = vT$ we have as before

$$y = a \sin\left(\frac{2\pi}{\lambda} vt + \alpha\right).$$

If $v$ is the velocity of the particle in the circle, then $vT = 2\pi a$. Hence if the velocity in the circle be the velocity of propagation of the wave, and if the time T be the period of the wave, then the circumference of the circle is equal to the length of the wave, while its radius represents the excursion of the molecule making simple harmonic vibrations.

The motion of the foot of the perpendicular is represented in Fig. 8, the variables being $y$ and $t$, and the co-ordinates $y$ and $\omega t$. Similarly the foot M of the perpendicular from P on OX executes the vibration

$$x = a \cos(\omega t + \alpha) = a \cos\left(\frac{2\pi t}{T} + \alpha\right).$$

**37. Algebraical Expression.**—Let O be a fixed origin and $y$ the displacement of P, at any instant, from its position of rest. Then if $x$ be the distance of P from O, measured along OX, the particles adjacent to P will in their displaced positions lie on a curve, as indicated in Fig. 9, and the equation of this curve at the instant under consideration may be written in the form $y = \phi(x)$.

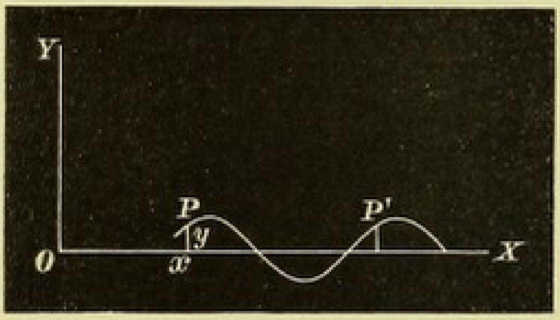


Fig. 9.

Now by the characteristic of wave motion this curve will be propagated forward with a velocity $v$, and the displacement of any particle P′ will be exactly the same as that of P at a time $t$ previously, provided the difference of the abscissæ of P′ and P is equal to $vt$, where $t$ is the time of propagation from P to P′. In fact the wave curve at P′ is merely that at P moved forward through a distance $vt$. But at P we have $y = \phi(x)$, therefore at P′ we have

$$y = \phi(x - vt),$$ [1]

for the $x$ of P is equal to the $x$ of P′ diminished by $vt$.

It is clear in itself that this equation represents a wave motion

[1] The form $y = a \sin\left\{\frac{2\pi}{\lambda}(vt - x) + A\right\}$ was tacitly assumed by Newton for the equation of motion of the air particles in sound propagation (*Principia*, lib. ii. prop. 47).

travelling with a velocity $v$, for $y$ remains unchanged if we replace $t$ by $t+\mathrm{T}$ and $x$ by its corresponding value $x+v\mathrm{T}$.

*Otherwise thus*:—If $v$ be the velocity of propagation, the displacements at any two points P and P′ will be the same at the times $t$ and $t'$ if

$$x'-x=v(t'-t) \quad \text{or } x'=x-vt+vt'.$$

Hence

$$y=\phi(x,\ t)=\phi(x',\ t')=\phi(x-vt+vt',\ t'),$$

which holds for all values of $t'$, and consequently for $t'=0$, in which case we have

$$y=\phi(vt-x).$$

If another disturbance be running in the opposite direction its equation will be $y=f(vt+x)$, so that if the two be superposed we will have

$$y=\phi(vt-x)+f(vt+x),$$

which represents the general state of disturbance in a finite stretched wire when we have both direct and reflected waves.

A particular solution of the equation $y=\phi(vt-x)$ is obviously

$$y=a\sin\left(\frac{vt-x}{\rho}\right).$$

The value of $y$ will be unaltered if $t$ be increased by $2\pi\rho/v$. The periodic time is therefore $\mathrm{T}=2\pi\rho/v$, but $\lambda=v\mathrm{T}=2\pi\rho$. Therefore [1]

$$y=a\sin\frac{2\pi}{\lambda}(vt-x).$$

The general form of $y$ is a periodic function. It has the same values at distances $x$, $x+\lambda$, $x+2\lambda$, etc., and at times $t$, $t+\mathrm{T}$, $t+2\mathrm{T}$, etc., and we know by Fourier's theorem (Thomson and Tait, *Nat. Phil.*) that any periodic function $f(z)$ of period $\lambda$ can be expanded as a series in the form

$$f(z)=\mathrm{A}_1\sin\left(\frac{2\pi z}{\lambda}+a_1\right)+\mathrm{A}_2\sin\left(\frac{4\pi z}{\lambda}+a_2\right)+\mathrm{A}_3\sin\left(\frac{6\pi z}{\lambda}+a_3\right)+\text{etc.}$$

Hence we may write

$$y=\mathrm{A}_1\sin\frac{2\pi}{\lambda}(vt-x_1)+\mathrm{A}_2\sin\frac{4\pi}{\lambda}(vt-x_2)+\text{etc.}$$

The complete vibration is therefore made up of a series of superposed simple harmonic vibrations of wave lengths $\lambda$, $\frac{1}{2}\lambda$, $\frac{1}{3}\lambda$, etc.

As we have reason to believe that waves of different periodic times

[1] Hence $\rho$ is the radius of a circle of circumference $\lambda$.

produce different impressions on the eye, just as notes of different pitch produce different effects on the ear, we may therefore restrict our attention to waves of a definite length, and take

$$y = a \sin \frac{2\pi}{\lambda}(vt - x) = a \sin \left(\frac{2\pi t}{T} - a\right)$$

Infinite train of waves.

or

$$y = a \sin (\omega t - a)$$

as the standard equation of the disturbance in our investigations concerning light of a definite wave length.

*Definition.*—The whole angle $2\pi(vt - x)/\lambda$, or $\omega t - a$, is called the *phase* of the vibration, and the angle $a$ is sometimes called its epoch.

**38. Colour and Frequency or Pitch.**—We have seen how a single wave or a continuous succession of waves (Fig. 6) may travel along a stretched cord. In a similar manner a continuous system of waves might be generated in the air or any other medium. Such a system of waves may or may not affect our sense of hearing. The phenomenon of sound is produced by a more or less rapid succession of waves, but there are major and minor limits to the rapidity of vibration, outside of which the ear fails to follow, and no sensation of sound is produced—that is, the sensibility of the ear is limited, and the pitch of the note must lie within certain limits in order that it may affect the sense of hearing. So a wave system in the ether may or may not affect our sense of sight. The sensations of light and colour are due to a very rapid succession of waves, but the sense of sight is limited in range and the ether vibration may be either too slow or too rapid to affect it. The limits in this case, however, are not nearly so widely separated as in the case of sound. The rapidity with which the violet waves succeed each other has been calculated to be less than twice as fast as that of the red, the latter making about four and the former about seven hundred millions of millions of vibrations per second. The sensibility of the eye is thus confined to much narrower limits than that of the ear, the interval between the red and violet being less than an octave, while the ear possesses a range of several octaves. Vibrations too slow to excite the eye are recognised by their thermal effects, and those which are too quick are generally detected by their chemical action. The former are most effective in heating our bodies, while the latter facilitate the prosecution of such arts as photography,[1] and are a very important factor in the growth of plants, etc.

Range limited.

[1] Captain Abney has succeeded in preparing photographic plates with bromide of silver which are capable of being decomposed not only by the violet end of the spectrum, but also by the red rays, and by rays of lower refrangibility which have wave lengths nearly three times that of the red (see "Bakerian Lecture," *Phil. Trans. Roy. Soc.*, 1881).

As the pitch (or musical colour) of a note is determined by the frequency of its vibrations, so it appears to be the frequency of the vibration in the luminiferous ether that determines the colour. For if a wave motion is propagated from one medium to another, the vibration frequency in the second medium ought to be the same as in the first, the vibration in the second being excited and forced by that in the first. This being admitted, it follows that if the velocity of propagation changes in passing from one medium to another, the wave length must change in the same proportion, in accordance with the equation $\lambda = v\mathrm{T}$. Now it has been observed that when light of a definite wave length passes from one medium to another, as from air to water, the colour of the light (say sodium light) remains unaltered, so that the natural inference is that the colour impression on the eye depends on the vibration frequency rather than on the absolute wave length.

We have now three methods of detecting solar radiation or wave motion in the ether. If the frequency of the vibration lies within certain limits, it affects our sight, and we call it light. Outside these limits the vibration may be too slow or too rapid to affect our eyes. In the former case we can detect the wave motion by its heating effects. It imparts its energy to material substances and to our own bodies, exciting vibrations in them and producing physical changes, and we term it radiant heat. It is not to be understood that it is only the waves too slow to affect our eyes that possess any heating power. This property is possessed by the luminous waves also, but generally in a smaller degree, so that those waves which affect our sense of sight may be called *luminous* heat waves, and those which do not may be termed *non-luminous*, or dark, heat waves. Both the ultra-violet and infra-red waves may be converted into luminous waves. The conversion of the former is termed *fluorescence* and of the latter *calorescence* (Arts. 289, 290).

Heat and light are consequently reduced to the same cause, viz. wave motion in the ether, and any ether wave may be, at the same time, a light wave and a heat wave. It is not that there are two classes of waves, heat waves and light waves, but that we have two senses by which we can detect the same waves. They affect our bodies with warmth and our eyes with light if their wave period lies within the range of our sensibility.[1]

[1] Evidently in Young's opinion the heat rays and the actinic rays differ from the light rays only in their wave length or period, for he states that ". . . we must therefore call light an influence capable of entering the eye, and affecting it with a sense of vision. A body from which this influence appears to originate is called a

The waves which are too rapid to affect our sense of sight are termed *actinic* waves or rays. Their extreme shortness probably qualifies them to operate between the molecules of bodies and become effective in producing chemical changes. Beyond this point we have no means of detecting the very rapid waves. It is otherwise, however, in the case of long waves.

Long waves.

We have seen that although the ether vibrations may be too slow to affect our sight (as air vibrations may be too slow to affect our sense of hearing), still we are able to detect their presence by their heating effect. But there is a limit even to this. The ether waves may be so slow that they will pass through our bodies without giving up any sensible portion of their energy. How then are we to detect these long ether waves? The recent brilliant experiments of Professor Hertz (see chap. xxi.) have placed the means in our hands. We can now work with ether waves of any length, from a few inches upwards. Heretofore we could detect only the radiant heat waves, which are excessively short, ranging from $\frac{1}{10000}$ to $\frac{1}{100000}$ part of an inch. The Hertzian waves are excited electrically and detected electrically. They may also be detected by their thermal effects. We have thus, as it were, acquired another sense, for we have now another means of investigating the ether, and connecting the various phenomena of nature.

**39. Average Kinetic Energy of a Vibrating Particle.**—Let a particle vibrate according to the equation

$$y = a \sin (\omega t - \alpha).$$

Then its velocity at any instant is given by

$$v = \frac{dy}{dt} = a\omega \cos (\omega t - \alpha),$$

and if $m$ be its mass its kinetic energy is $\frac{1}{2}mv^2$. Now the foregoing expression shows that $v$ varies from zero to $a\omega$; accordingly the mean energy during a complete vibration will be

$$\frac{1}{T}\int_0^T \tfrac{1}{2}mv^2 dt = \frac{ma^2\omega^2}{4T}\int_0^T 2\cos^2(\omega t - \alpha)dt = \frac{ma^2\omega^2}{4T}\int_0^T \left\{1 + \cos 2(\omega t - \alpha)\right\}dt$$
$$= \frac{ma^2\omega^2}{4T}\left\{t + \frac{1}{2\omega}\sin 2(\omega t - \alpha)\right\} = \tfrac{1}{4}ma^2\omega^2,$$

luminous body. We do not include in this definition of the term light the invisible influences which occasion heat only, or blacken the salts of silver, although they both appear to differ from light in no other respects than as one kind of light differs from another, and they might probably have served the purpose of light if our organs had been differently constituted" (Young, "On the Theory of Optics," *Lectures on Natural Philosophy*).

where $\omega = 2\pi/T$. Hence the average energy is

$$= \frac{ma^2\omega^2}{4} = \frac{m\pi^2a^2}{T^2}.$$

*Cor.*—The mean energy $= \frac{1}{4}mv^2 = \frac{1}{2}\cdot\frac{1}{2}mv^2$ where $v$ is the greatest velocity of the molecule. Therefore the mean kinetic energy is half the maximum kinetic energy.

**40. Intensity of Illumination.**—We must now define what we mean by, and how we measure, the *intensity* of light. Lights generally differ in two respects, viz. intensity and colour. In photometry we say that two lights are equally intense if at the same distance away they produce the same illumination on a screen. If one candle produces a certain illumination we say that two such candles together produce twice as much illumination, or that the *intensity* of illumination is twice as great. Similarly the intensity of illumination is directly proportional to the number of candles, provided they are all of the same power.[1] Now the average energy of the vibration due to any source is proportional at each point to the square of the amplitude, and the average energy due to any number of sources is proportional to the square of the amplitude of the resultant vibration, and this over any region is the sum of the average energies due to the sources separately.

This has led to two methods of estimating the intensity of the illumination in terms of the energy of the vibrating medium. In the first place, we may take the intensity of illumination as measured by the energy per unit volume of the medium, and in this case if $m$ stands for the mass per unit volume the intensity by the foregoing article will be measured by the formula[2]

$$I = \frac{m\omega^2a^2}{4} = \frac{m\pi^2a^2}{T^2}.$$

Energy contained.

On the other hand, we may measure the intensity by the quantity of energy transmitted per unit time per unit area across a plane perpendicular to the direction of propagation. In this case the velocity of propagation will come into account, and the intensity (for a plane wave) will be measured by the product of the velocity and the energy per unit volume—that is, by the formula

$$I = \frac{m\omega^2a^2}{4}v = \frac{m\pi^2a^2\lambda}{T^3}.$$

Energy transmitted.

[1] This may be regarded as the definition of an illumination $n$ times as intense as a given standard. That the intensity varies inversely as the square of the distance from the source follows as an experimental fact.

[2] In this expression we have taken into account only the kinetic energy. In the case of a vibrating elastic solid the whole energy is half kinetic and half potential, so that if the whole energy be taken as the measure of I, the expressions in the text ought to be doubled.

In practice, however, we deal only with the relative intensities, and it follows from either system of measurement that if the periods of two vibrations be the same, then their intensities are in the ratio of the squares of their amplitudes, or

$$\frac{I}{I'} = \frac{a^2}{a'^2}.$$

The relative intensities of two sounds of the same pitch, or of two lights of the same colour—that is, of the same period or wave length—are consequently compared by the squares of their amplitudes of vibration. But we cannot so easily compare the intensities of two lights of different colour, or two sounds of different pitch, for they produce dissimilar impressions, and the time (T) of vibration enters the expression for the mean energy of the motion. In the estimation of the relative intensities of different sources of light (which forms the subject of photometry) great difficulty is encountered in the fact that different sources of light are in general differently coloured.[1]

**41. The Intensity varies inversely as the Square of the Distance.**—Let us now consider a luminous point as the source of a system of spherical waves diverging from it as centre. Consider any one of these waves of radius $r$. This wave travels outward, developing itself and increasing its radius with the velocity of light. It carries its energy with it, and after a time it is a sphere of radius $r'$. Now if I denotes the energy per unit time transmitted across unit area of the first wave surface, and I′ the energy per unit area of the second, we have, since the energy of the first is handed on to the second, the whole energy transmitted per unit time across the wave $= 4\pi r^2 I$, and also the whole energy transmitted $= 4\pi r'^2 I'$.

Consequently

$$\frac{I}{I'} = \frac{r'^2}{r^2}, \quad \text{or } Ir^2 = I'r'^2 = \text{constant},$$

or the intensities of the illumination at different distances from a luminous origin are inversely as the squares of the distances, a relation which is verified by experiment.

*Cor.*—Since the intensity is proportional to the square of the amplitude, it follows that the amplitude of the vibration is inversely as the distance from the origin, or

$$ar = a'r' = \text{constant}.$$

---

[1] On the comparison of the relative intensities of the different colour throughout the spectrum see *Colour Measurement and Mixture*, by Captain Abney.

**42. The Principle of Superposition.**—When two or more separate disturbances are simultaneously impressed on the same element of a medium, the effect may be very complex, but in the case of light the displacements are supposed to be such that the composition of them by direct superposition is allowable, for calculations made on this assumption agree with the observed facts, at least to that degree of accuracy reached by experiment. On this principle is based the whole doctrine of interference discovered by Young in 1801. It follows from it that any number of separate disturbances may be propagated *through* one another in the same portion of the medium, each emerging from that portion as if it had not been encountered by the others. Thus rays of light from objects all around us cross each other's paths in all sorts of ways, but each travels on as if the others did not exist. Each portion of the ether is traversed by streams of light from a multitude of different sources, which are simultaneously propagated through it. Hence we may conclude that at any instant the disturbance at any point of the ether is that due to the superposition of all the disturbances which reach it at that instant from the various parts of the surrounding medium. This is, in a generalised form, the principle stated by Huygens in 1678. It asserts that the displacement caused by any source of small vibrations is the same, whether it acts upon the medium alone or in conjunction with other sources, provided the displacements considered are very small. The combined effect of several sources is therefore the geometric resultant of the displacements which would be produced by them acting separately.

The nature of this principle may be made more clear by a simple example. Let a pendulum receive an impulse in any vertical plane passing through the point of suspension, causing it to vibrate in that plane. Now when it is at the lowest point of the arc of vibration, let a second impulse be given horizontally in a plane perpendicular to that in which it already vibrates. This impulse, if it had acted on the pendulum at rest, would have caused it to vibrate in the vertical plane of the impulse and through an arc depending on the magnitude of it. Now it is found on trial that the distance of the bob of the pendulum from either of these vertical planes is the same at any instant as if the other vibration did not exist, so that each vibration subsists independently of the other, although the resultant motion is a compound elliptic vibration.

The two vibrations are here taken in separate planes in order that their coexistence may be more easily recognised. When the vibrations are in the same plane the resultant vibration is also in that plane, and its amplitude, by the principle of superposition, is the sum of the amplitudes of the constituents when their directions conspire, and their difference when they are opposed.

**43. To compound two Simple Vibrations.**—Let a particle be simultaneously impressed by the periodic displacements represented by the equations [1]

$$y_1 = a_1 \sin(\omega t - \alpha_1) \quad \text{and} \quad y_2 = a_2 \sin(\omega t - \alpha_2),$$

which agree in periodic time and wave length ($\omega = 2\pi/T$), but differ in phase and amplitude. The resultant displacement at any time $t$ is

$$\begin{aligned} y = y_1 + y_2 &= a_1 \sin(\omega t - \alpha_1) + a_2 \sin(\omega t - \alpha_2), \\ &= \sin \omega t(a_1 \cos \alpha_1 + a_2 \cos \alpha_2) - \cos \omega t(a_1 \sin \alpha_1 + a_2 \sin \alpha_2), \\ &= A \sin(\omega t - \alpha), \end{aligned}$$

if
$$A \cos \alpha = a_1 \cos \alpha_1 + a_2 \cos \alpha_2$$
and
$$A \sin \alpha = a_1 \sin \alpha_1 + a_2 \sin \alpha_2.$$

Therefore by squaring and adding we find

$$A^2 = a_1^2 + a_2^2 + 2a_1 a_2 \cos(\alpha_1 - \alpha_2),$$

and by division [2]

$$\tan \alpha = \frac{a_1 \sin \alpha_1 + a_2 \sin \alpha_2}{a_1 \cos \alpha_1 + a_2 \cos \alpha_2}.$$

---

[1] Fresnel, *Œuvres complètes*, tom. i. pp. 288-293 ; *Mémoire couronné sur la Diffraction*, §§ 37-42.

[2] These results may be obtained geometrically as follows:—Let the parallelogram OABC revolve round the vertex O with uniform angular velocity. Then since A, B, C describe circles round O with a uniform angular velocity, it follows by Ex. 2, p. 37, that the feet P, Q, R of the perpendiculars from A, B, C on any right line OX execute simple harmonic vibrations along the line OX. But since AC is equal and parallel to OB, it follows that PR = OQ, and therefore OR = OP + OQ. Hence the displacement of R at any instant is equal to the sum of the displacements of P and Q ; or, in other words, the motion of R is the resultant of the motions of P and Q superposed. Now R executes a simple harmonic vibration of amplitude OC and phase XOC—that is, if the amplitudes of the vibrations of P and Q be represented by the lines OA and OB, while the phases of the same vibrations are represented by the angles OA and OB make with a fixed line OX, then the amplitude of the resultant of the two will be denoted by OC, the diagonal of the parallelogram on OA and OB, and the phase of the resultant will be represented by the angle OC makes with OX. The analytical method in the text is given in order that the student may become armed at once with a direct and powerful instrument of attack. [Compare Art. 36, Ex. 2. Figure 10 illustrates the cosine formulæ. The sine formulæ may, however, be deduced from the figure by supposing the phase measured clock-wise from the vertical OY.]

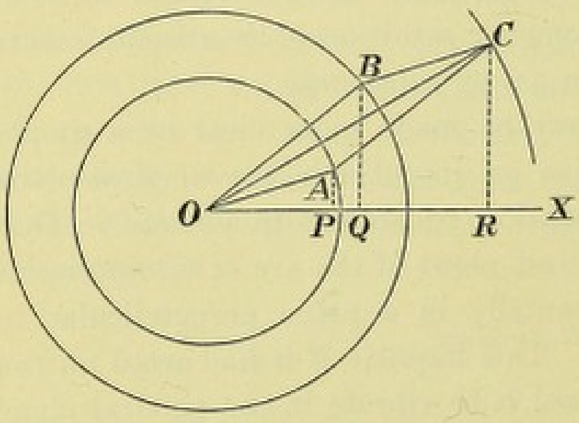


Fig. 10.

This determines the amplitude A and the phase $\omega t - \alpha$ of the resultant vibration. If we denote the difference of phase of the component vibrations by $\delta$ we have $\delta = \alpha_1 - \alpha_2$, and the resultant amplitude is given by the equation

$$A^2 = a_1^2 + a_2^2 + 2a_1a_2 \cos \delta.$$

*Cor.* 1.—If $\delta = 0$ or $2n\pi$—that is, if one wave is retarded on the other by any number of wave lengths—we have $\cos \delta = 1$, and

$$A^2 = (a_1 + a_2)^2.$$

*Cor.* 2.—If $\delta = \pi$ or $(2n + 1)\pi$, one wave is retarded an odd number of half wave lengths on the other, and

$$A^2 = (a_1 - a_2)^2.$$

If in addition $a_1 = a_2$, we have $A = 0$, and the waves in this case destroy each other.

*Cor.* 3.—If $\delta = \frac{1}{2}\pi$ or $(n + \frac{1}{2})\pi$, one wave is retarded a quarter wave on the other, or $(n + \frac{1}{2})\frac{1}{2}\lambda$ and

$$A^2 = a_1^2 + a_2^2.$$

**44. Distribution of the Energy.**—In the case of a medium simultaneously disturbed by two sources of similar vibrations, we have seen that the amplitude of the resultant vibration at any point P is determined by the equation $A^2 = a_1^2 + a_2^2 + 2a_1a_2 \cos \delta$, where $\delta$ is the phase difference of the vibrations reaching P, and depends on the difference of its distances from the sources. If this difference is to remain fixed, then P may lie anywhere on a hyperboloid of revolution, having the sources for foci. The whole space around the sources may be divided up by a system of such surfaces, each determined by a value of $\delta$ and a corresponding value of A. The average energy of the vibration of the elements of the medium at any one of these surfaces is measured by $A^2$, which, from surface to surface, varies from a maximum $(a_1 + a_2)^2$, for $\delta = 2n\pi$, to $(a_1 - a_2)^2$, for $\delta = (2n + 1)\pi$. The average value, however, of $A^2$ throughout the whole medium will be

$$a_1^2 + a_2^2,$$

or the sum of the average energies of the vibrations excited by the sources acting separately.

Thus when a medium is disturbed simultaneously by two similar sources $O_1$ and $O_2$, the amplitude of the resultant vibration at any point is the sum of the amplitudes of the vibrations which the sources would excite if each acted separately only when the vibrations at this point are in the same phase. The amplitude at any other point will depend on the difference of phase, so that in some tracts of the medium there is a large amount of energy, while in others there is very little. The interfering action of the two sources causes no destruction of energy, but merely a redistribution of it in the medium around them.

For example, if $a_1 = a_2$, that is, if the sources be of equal power, then at some parts of the medium the energy will be $(2a)^2$, or quadrupled, while at others it will be zero, but the average value throughout the medium will be $2a^2$, viz. the sum of the average energies of the two equal sources.

### *Example*

Any number of waves of a given type compound into one of the same type. For if

$$y = y_1 + y_2 + \ldots + y_n = \Sigma a_1 \sin(\omega t - \alpha_1),$$
$$= Q \sin \omega t - P \cos \omega t = A \sin(\omega t - \alpha).$$

Where $$P = \Sigma a_1 \sin \alpha_1,\quad Q = \Sigma a_1 \cos \alpha_1,$$

$$\tan \alpha = \frac{\Sigma a_1 \sin \alpha_1}{\Sigma a_1 \cos \alpha_1} = \frac{P}{Q},$$

and $$A^2 = \Sigma a_1^2 + 2\Sigma a_1 a_2 \cos(\alpha_1 - \alpha_2) = P^2 + Q^2.$$

If the epoch angles $\alpha_1$ and $\alpha_2$, etc., vary irregularly, the chance is that the term involving $\cos(\alpha_1 - \alpha_2)$ will be as much positive as negative,[1] and will disappear from the result, so that the average value of $A^2$ will be $\Sigma a^2$, or

average intensity = sum of component intensities.

**45. Graphic Representations of the Resultant of a System of Vibrations.**—From the equation $A^2 = a_1^2 + a_2^2 + 2a_1 a_2 \cos \delta$ it appears that the resultant amplitude A is the diagonal of a parallelogram of which the sides are $a_1$ and $a_2$, and $\delta$ the angle between them.

Again the equation for tan $\alpha$ may be written in the form

$$a_1 \sin(\alpha - \alpha_1) + a_2 \sin(\alpha - \alpha_2) = 0,$$

from which it follows that if we draw a line making an angle $\alpha$ with the diagonal A, it will make angles $\alpha_1$ and $\alpha_2$ with the sides $a_1$ and $a_2$

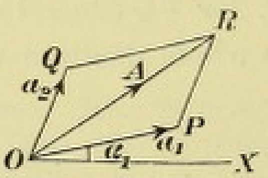



Fig. 11.—Resultant of two similar Vibrations.

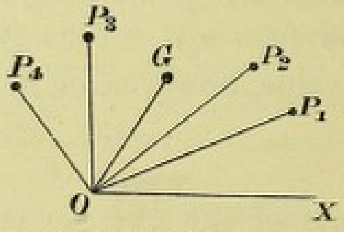



Fig. 12.—Resultant of a system of similar Vibrations.

of the parallelogram. Thus we have the following construction. Take any fixed line OX (Fig. 11) and draw OP equal to $a_1$, and making an

[1] On this point see *Ency. Brit.* ninth edition, art. "Wave Theory," by Lord Rayleigh. Also *Phil. Mag.*, Aug. 1880.

angle $a_1$ with OX, similarly draw OQ equal to $a_2$, and making $QOX = a_2$, then if we complete the parallelogram OPRQ the diagonal OR is equal to A, and the angle ROX which it makes with OX is equal to $a$, while POQ is equal to the difference of phase $\delta$ or $a_1 - a_2$. Or if POX and QOX represent the phases $\phi_1$ and $\phi_2$ of the component vibrations, ROX will be the phase $\phi$ of the resultant vibration. The amplitudes, therefore, compound like forces.

To compound several vibrations by this method we have only to draw $OP_1$, $OP_2$, $OP_3$, etc. (Fig. 12), from O, equal to $a_1$, $a_2$, $a_3$, etc., the amplitudes of the vibrations, and making angles $P_1OX$, $P_2OX$, etc., equal to phases $\phi_1$, $\phi_2$, etc.; then, as in the case of a system of forces (Minchin's *Statics*), the resultant amplitude is equal to $n$.OG, if there are $n$ amplitudes, and if G is the centre of mean position of the points $P_1$, $P_2$, $P_3$, etc.

It may be more instructive, however, to employ the method of the polygon of forces. Thus draw $OP_1$ (Fig. 13) equal to $a_1$, and making an angle $\phi_1$ with OX, and from $P_1$ draw $P_1P_2 = a_2$, and making $\phi_2$ with OX. Then $OP_2$ represents the resultant of the vibrations $a_1$ and $a_2$. Similarly for a third draw $P_2P_3 = a_3$, and making an angle $\phi_3$ with OX, and so on for any number. The line joining O to $P_n$, the extremity of the line last drawn, represents in magnitude the amplitude, and its direction makes an angle with OX equal to the phase of the resultant vibration.

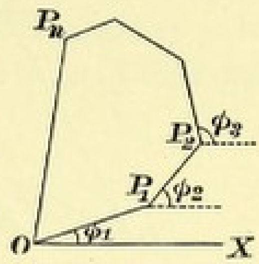


Fig. 13.—The Vibration Polygon.

These graphic methods we shall find of great utility in the theory of diffraction.

*Cor.* 1.—When the system of vibrations to be compounded forms a series in which the phase difference of any two consecutive terms is infinitesimal, but varying continuously from term to term, then the angle between any two consecutive sides of the vibration polygon will be infinitesimal. The polygon will consequently become a continuous curve, as shown in Fig. 14, such that the element of length at any point M of the curve represents the amplitude of a corresponding constituent vibration, and the angle which tangent at this point makes with OX represents the phase $\phi$ of the same vibration. The line joining O to M represents the resultant of the system of vibrations between O and M; the length of the line OM representing its amplitude, and the angle MOX its phase. If P be a point on the curve such that the tangent at P is parallel to OX, then the arc OMP represents a system of vibrations varying continuously in phase from zero to $\pi$—that is, over half a period—and OP

Vibration curve.

represents the amplitude, and XOP the phase of the resultant of the system.

*Cor.* 2.—If the system of vibrations forms a series of equal amplitude and equicrescent phase, then the sides of the vibration polygon will be equal and equally inclined to each other. In other words, the vibration polygon will be regular and inscribed in a circle. Hence, when the number of terms is very great and the constant phase difference very small, the polygon will, in the limit, coincide with its circumscribing circle, and it follows that the resultant of an infinite number of vibrations of equal amplitude, and varying uniformly in phase, may be represented in amplitude and phase by the chord of a circular arc. The general curve of Fig. 14 becomes in this case a circle, as shown in Fig. 15. The arc OM of this circle represents a system of vibrations of equal amplitude and phase varying uniformly

Vibration circle.

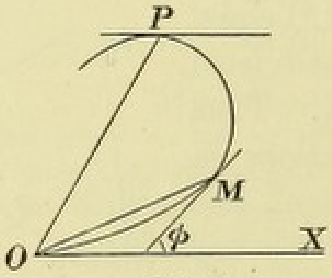



Fig. 14.

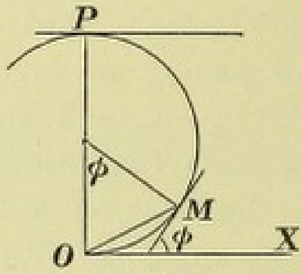



Fig. 15.

from zero to $\phi$. The chord OM represents the resultant, and the amplitude is therefore

$$OM = 2r \sin\tfrac{1}{2}\phi,$$

while the phase of the resultant is

$$MOX = \tfrac{1}{2}\phi.$$

That is, the phase of the resultant is the arithmetic mean of the initial and final constituents of the series, and it is consequently the same as the phase of the constituent corresponding to the middle point of the arc OM; for, in the case of a circle, the tangent at the middle point of the arc is parallel to the chord of the arc.

The diameter OP of the circle represents the resultant of a system of vibrations of equal amplitude, and phase varying uniformly between the limits 0 and $\pi$. The phase of this resultant is $\frac{1}{2}\pi$, and it is therefore a quarter period behind the first constituent of the series, or the same as the middle term of the series. If the amplitude of each constituent of the series be $a$, and if there be $n$ terms in the series, then if the series be represented by any arc OM of a circle, we have in the limit for the length of the arc

$$\text{Arc } OM = na.$$

Hence, if $r$ be the radius of the circle and $\phi$ the difference of phase of the first and last terms of the series, we have

$$r\phi = na.$$

Consequently the amplitude of the resultant is

$$\mathrm{OM} = 2r \sin \tfrac{1}{2}\phi = na \frac{\sin \frac{1}{2}\phi}{\frac{1}{2}\phi}.$$

If the phase varies from 0° to $\pi$, the arc will be a semicircle and the amplitude of the resultant will be $2r$—that is,

$$\frac{2na}{\pi}.$$

**46. Waves of Different Lengths.**—We have good reason to believe that the velocity of propagation in air is very nearly the same for light waves of every length, just as we know that sound waves of different lengths travel with the same velocity. Still whether $v$ be constant or not it is generally impossible for two vibrations of different periods such as

$$y_1 = a_1 \sin(\omega_1 t + \alpha_1), \quad \text{and } y_2 = a_2 \sin(\omega_2 t + \alpha_2)$$

to completely destroy each other. We have seen that when $\omega_1 = \omega_2 = \omega$, any number of such waves combine to form a resultant wave A $\sin(\omega t - \alpha)$ of the same period as the constituents, and containing no trace of their distinction. The two waves given above cannot, however, be so compounded unless their periodic times be the same—that is, unless we have $v_1/\lambda_1 = v_2/\lambda_2$. The consideration, therefore, of waves of different lengths may be kept perfectly separate, for their ultimate effect will be the superposition of their separate effects, each possessing its own individuality.

This point is illustrated by sounding two notes of different pitches together (for example, a note and its fifth); or by the mixture of two colours.

**47. To compound two Rectangular Vibrations.**—If a particle be subjected simultaneously to two simple harmonic vibrations of the same period, but in perpendicular directions, the resultant vibration is in general elliptic.

Let the superposed vibrations along OX and OY (Fig. 16) be given by the equations

$$x = a \sin\theta,$$
$$y = b \sin(\theta + \delta)$$

when $\theta$ for brevity denotes the phase of one of the vibrations and $\delta$

their difference of phase. Now $\theta$ is a variable containing the time, and to eliminate it we have

$$y = b\,(\sin\theta\cos\delta + \cos\theta\sin\delta),$$

and
$$\sin\theta = x/a,$$

therefore
$$\frac{y}{b} = \frac{x}{a}\cos\delta + \sin\delta\sqrt{1 - \frac{x^2}{a^2}}.$$

Hence
$$\left(\frac{y}{b} - \frac{x}{a}\cos\delta\right)^2 = \sin^2\delta\left(1 - \frac{x^2}{a^2}\right),$$

or finally
$$\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{2xy}{ab}\cos\delta = \sin^2\delta.$$

This equation denotes an ellipse having its centre at O, which is consequently the curve described by the particle under the simultaneous

Elliptic Vibrations.

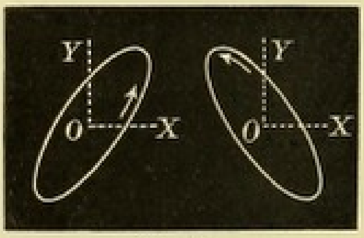



Fig. 16.

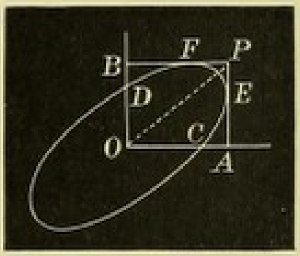



Fig. 17.

action of two rectangular vibrations of the same periodic time, of amplitudes $a$ and $b$, and difference of phase $\delta$.

If $\delta$ be increased by $\pi$ the equation becomes

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{2xy}{ab}\cos\delta = \sin^2\delta.$$

The greatest values of $x$ and $y$ are $a$ and $b$ respectively, hence if tangents AE and BF be drawn parallel to the axes of reference, we have (Fig. 17)

$$\text{OA} = a, \quad \text{and OB} = b.$$

Also by putting $x = 0$ and $y = 0$ successively in the equation of the curve we find

$$\text{OC} = a\sin\delta, \quad \text{and OD} = b\sin\delta.$$

Hence

$$\sin\delta = \frac{\text{OC}}{\text{OA}} = \frac{\text{OD}}{\text{OB}}.$$

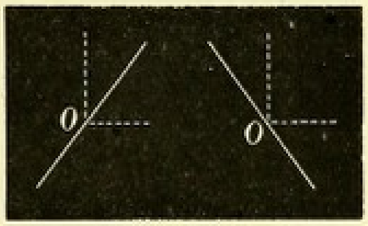
Again if we replace $x$ by its maximum value $a$, we find $y = b\cos\delta$ for the ordinate of the point E; and if we replace $y$ by its maximum value $b$ we find $x = a\cos\delta$ for the abscissa of the point F, therefore

$$\cos\delta = \frac{AE}{AP} = \frac{BF}{BP}.$$

*Cor.* 1.—If the difference of phase $\delta$ be zero or $2n\pi$ the equation of the resultant vibration becomes

$$\frac{x}{a} - \frac{y}{b} = 0.$$

This represents a right line passing through O (Fig. 18) and making an angle with the axis of $y$, the tangent of which is $a/b$.




Figs. 18 and 19.—Phase Difference a Multiple of $\pi$.

*Cor.* 2.—If the difference of phase be $\pi$ or $(2n+1)\pi$, which corresponds to a retardation $\frac{1}{2}\lambda$ or $(2n+1)\frac{\lambda}{2}$, the resultant vibration is represented by

$$\frac{x}{a} + \frac{y}{b} = 0.$$

This is the equation of a right line passing through O (Fig. 19) and inclined to the axis of $y$ on the other side at an angle whose tangent is $a/b$.

Thus we see that if two simple rectangular vibrations be compounded, the resultant vibration is a simple rectilinear vibration if they differ in phase by any number of times $2\pi$ (or a retardation of any number of times $\lambda$), and if a difference of phase equal to any odd multiple of $\pi$ or a retardation of any odd multiple of half a wave length be introduced, the resultant vibration is still rectilinear, but its direction is turned through an angle $2\theta$, if $\theta$ be the angle the first vibration makes with either axis.

This remark is of importance in the theory of polarised light.

*Cor.* 3.—If $\delta = \frac{1}{2}\pi$, or any odd multiple of $\frac{1}{2}\pi$, that is, if there is a retardation of $\frac{1}{4}\lambda$, we have

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1.$$

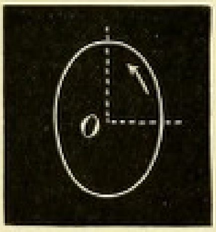



Fig. 20.—Phase Difference a Right Angle.

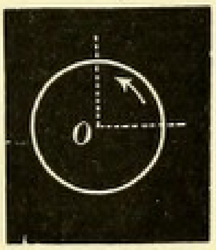



Fig. 21.—Circular Vibration.

In this case the axes of the elliptic vibration coincide with the directions of the component vibrations (Fig. 20).

In the first cases the two rectangular vibrations pass through their middle points at the same instant, and in this case one passes through its middle point when the other is at its extremity.

If, in addition, $a = b$, we have

$$x^2 + y^2 = a^2,$$

and the resultant vibration is circular (Fig. 21).

Thus two rectangular vibrations of equal amplitude and periodic time will compound into a circular vibration if they differ in phase by any odd multiple of $\frac{1}{2}\pi$, or if one be retarded on the other by any odd multiple of $\frac{1}{4}\lambda$.

**48. Vibration of Permanent Type—Polarised Light.**—It has been remarked already that the phenomena of interference lead us to believe that light is propagated as a periodic disturbance or a periodic change of some sort in the condition of a medium. This periodic change is referred to as the vibration; but as to the nature of the change, whether it is simply a periodic displacement of the elements of the medium, such as occurs in a vibrating elastic solid, or as to whether it is something of quite a different nature must remain a matter of speculation. In representing the vibration by an equation of the form

$$y = a \sin (\omega t + \alpha)$$

no particular assumption need be made as to what it is that $y$ represents, except that it is the change of condition at the time $t$. It is perhaps simplest to regard $y$ as a displacement in the medium such as occurs in ordinary wave motion. This is the ordinary notion, but there is no advantage in imposing such a limitation on the nature of the disturbance, except for the purposes of distinct conception in the mind.

In representing a periodic change by a simple equation of the foregoing form it should be remembered, however, that this equation embraces an infinite succession of similar changes—that is, an infinite train of waves of invariable type. Light is said to be polarised when the type of the vibration is maintained invariable. If the ether vibration consists in a periodic displacement, all the elements of the vibrating ether in the wave front describe similar, and similarly situated, curves which remain *permanently* the same. Thus if all the elements describe similar and similarly situated ellipses for any time the light is said to be elliptically polarised during that time. If they describe circles it is circularly polarised, and if they describe parallel rectilinear paths it is called plane-polarised light. The essential feature of polarised light is then that all the elements of ether in any wave front continue to describe exactly the same kind of orbits, or the nature of the disturbance remains permanently the same.

## *Examples*

1. Any number of simple harmonic vibrations in different directions, differing in phase but having the same periodic time, compound into an elliptic vibration.

2. Any simple harmonic vibration is equivalent to two opposite circular vibrations.

[For $x=a\cos\omega t,\ y=a\sin\omega t,$

is a circular vibration, viz. $x^2+y^2=a^2$, so also, changing the sign of $\omega$ we have

$$x=a\cos\omega t,\ y=-a\sin\omega t$$

as a circular vibration in the opposite direction. Adding the $x$-components, and also the $y$-components, we find for the resultant vibration

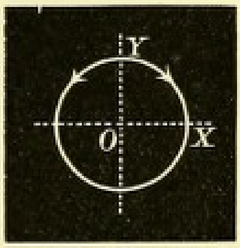


Fig. 22.

$$x=2a\cos\omega t,\ y=0.$$

Consequently the two opposite circular vibrations compound into a simple rectilinear vibration of double the amplitude.

Conversely, if we start with the simple vibration $x=a\cos\omega t$, we may write it in the form

$$(1)\ \ x=\tfrac{1}{2}a\cos\omega t,\ y=\ \ \tfrac{1}{2}a\sin\omega t.$$
$$(2)\ \ x=\tfrac{1}{2}a\cos\omega t,\ y=-\tfrac{1}{2}a\sin\omega t.$$

The first pair represent a circular vibration of amplitude $\frac{1}{2}a$, and the second pair represent an opposite circular vibration of amplitude $\frac{1}{2}a$.

This result is of importance in the theory of circularly polarised light.

An inspection of the figure shows at once that a particle moving with simple harmonic motion along a line OY may be supposed actuated by two equal and opposite circular vibrations of half the amplitude, for the velocities perpendicular to this line destroy each other, while the velocity along it is doubled.]

3. Two simple harmonic vibrations, in rectangular directions, compound into a parabolic vibration if the periodic time and phase of one is double that of the other.

[The equations of the component vibrations are

$$x=a\cos 2(\omega t+\alpha)=a\cos 2\theta$$
$$y=b\cos(\omega t+\alpha)=b\cos\theta,$$

therefore eliminating $\theta$, we find

$$\frac{2y^2}{b^2}=\frac{x}{a}+1.$$

If the phase of one is not double that of the other, the resultant vibration is of the fourth degree.

The above vibration is that obtained by compounding a note and its octave. A further exercise is that of compounding a note and its fifth (periods in the ratio 2 : 3), or any note and another at an octave + fifth interval from it (periods in ratio 1 : 3). Diagrams of these cases will be found in treatises on Sound under the head "Lissajous' figures."]

4. If a system of vibrations differing in amplitude and wave length (or periodic time) be superposed, we have for resultant

$$y=a_1\sin 2\pi\left(\frac{t}{T_1}+\alpha_1\right)+a_2\sin 2\pi\left(\frac{t}{T_2}+\alpha_2\right)+\ .\ .\ .\ \text{etc.},$$

which may be written in the form

$$y = a_1 \sin 2\pi\left(\frac{m_1 t}{\mathrm{T}} + \alpha_1\right) + a_2 \sin 2\pi\left(\frac{m_2 t}{\mathrm{T}} + \alpha_2\right) + \ . \ . \ .$$

where $\mathrm{T} = m_1\mathrm{T}_1 = m_2\mathrm{T}_2 = $ etc., $m_1$, $m_2$, etc., being integers. This shows that $y$ is periodic, for it remains unaltered when for $t$ we substitute $t + \mathrm{T}$. The periodic time of the resultant vibration is therefore T, the L.C.M. of $\mathrm{T}_1$, $\mathrm{T}_2$, $\mathrm{T}_3$, etc.

5. The components of an elliptic vibration are

$$x = a \sin \omega t \quad \text{and} \quad y = b \sin(\omega t + \delta),$$

find the direction of the axes of the ellipse.

[The directions are given by $\tan 2\phi = \dfrac{2ab}{a^2 - b^2}\cos\delta$ where $\tan\phi = \dfrac{y}{x}$.]

6. Two elliptic vibrations, given by the equations

$$\left.\begin{array}{l} x_1 = a_1 \sin(\omega t + \alpha_1) \\ y_1 = b_1 \sin(\omega t + \beta_1) \end{array}\right\} \quad \left.\begin{array}{l} x_2 = a_2 \sin(\omega t + \alpha_2) \\ y_2 = b_2 \sin(\omega t + \beta_2) \end{array}\right\}$$

are superposed, find the resultant vibration.

[If the components of the resultant vibration be

$$x = a \sin(\omega t + \alpha) \quad \text{and} \quad y = b \sin(\omega t + \beta),$$

we have

$$a^2 = a_1^2 + a_2^2 + 2a_1a_2\cos(\alpha_1 - \alpha_2)$$
$$b^2 = b_1^2 + b_2^2 + 2b_1b_2\cos(\beta_1 - \beta_2)$$

$$\tan\alpha = \frac{a_1\sin\alpha_1 + a_2\sin\alpha_2}{a_1\cos\alpha_1 + a_2\cos\alpha_2}, \quad \tan\beta = \frac{b_1\sin\beta_1 + b_2\sin\beta_2}{b_1\cos\beta_1 + b_2\cos\beta_2}.\text{]}$$

7. The elliptic vibration

$$x = a \sin \omega t, \quad y = b \cos \omega t$$

is equivalent to the two superposed circular vibrations

$$\left.\begin{array}{l} x_1 = \frac{1}{2}(a+b)\sin\omega t \\ y_1 = \frac{1}{2}(a+b)\cos\omega t \end{array}\right\} \quad \left.\begin{array}{l} x_2 = \frac{1}{2}(a-b)\sin\omega t \\ y_2 = -\frac{1}{2}(a-b)\cos\omega t \end{array}\right\},$$

consequently the elliptic vibration may be regarded as the resultant of two oppositely directed circular vibrations of amplitudes, $\frac{1}{2}(a+b)$ and $\frac{1}{2}(a-b)$ respectively.

8. *System of Discrete Particles—Minor Limit to the Periodic Time.*—A system of pellets, each of mass $m$, is attached to a weightless string at equal intervals $a$. The string is stretched with tension T along a smooth horizontal plane, so that the force of gravity does not enter the question of the horizontal tranverse disturbances, discuss the motion (P. G. Tait, *Ency. Brit.* art. "Waves").

Let the transverse displacement of the $n$th pellet be $y_n$ at the time $t$. The equation of its motion is

$$m\frac{d^2y_n}{dt^2} = \mathrm{T}\frac{y_{n+1} - y_n}{a} - \mathrm{T}\frac{y_n - y_{n-1}}{a}.$$ [1]

---

[1] If $y_n = f(x)$, then $y_{n+1} = f(x+a) = e^{a\frac{d}{dx}}f(x)$, consequently we have $y_{n+1} = e^{a\frac{d}{dx}}y_n$ and $y_{n-1} = e^{-a\frac{d}{dx}}y_n$. Hence $m\dfrac{d^2y_n}{dt^2} = \dfrac{\mathrm{T}}{a}\left(e^{a\frac{d}{dx}} - 2 + e^{-a\frac{d}{dx}}\right)y_n$, which reduces to $\mathrm{T}a\dfrac{d^2y}{dx^2}$ when $a$ is very small, so that we recover the equation of vibration of a continuous cord.

Again, if $y_n = \text{A}\cos(\omega t - kx)$ where $x = na$, we have, from the above equation,

$$-m\omega^2\cos(\omega t - kx) = \frac{\text{T}}{a}[\cos\{\omega t - k(x+a)\} - 2\cos(\omega t - kx) + \cos\{\omega t - k(x-a)\}]$$

$$= -\frac{2\text{T}}{a}\cos(\omega t - kx)(1 - \cos ka).$$

Hence
$$m\omega^2 = \frac{4\text{T}}{a}\sin^2\tfrac{1}{2}ka.$$

The greatest possible value of $\omega$ is consequently $2\sqrt{\text{T}/ma}$—that is, we must have $\omega < 2v/a$, if $v$ be the velocity of a disturbance in a cord under the same tension and of the same mass per unit length. The time of oscillation of a pellet is $2\pi/\omega$ and is $> \pi a/v$ or $\pi\sqrt{am/\text{T}}$.

This question is closely connected with Stokes's theory of fluorescence. For if a disturbing force of shorter period than the limit given above be applied continuously to one of the pellets, there will be an accumulation of energy in its neighbourhood ; and this energy, if we suppose the disturbing force to cease, will be transmitted throughout the system by vibrations of equal or greater period than the limiting value above, corresponding to light of lower refrangibility than the incident, but having a definite superior limit of refrangibility.

9. *On the Group Velocity of a Train of Waves.*—When a group of waves advances into still water, the velocity of the group is less than that of the individual waves which constitute it. The individual waves appear to advance through the group, dying away as they approach its anterior limit. Sir G. G. Stokes first explained this by regarding the group as formed by the superposition of two infinite trains of waves, of equal amplitude, and of nearly equal wave lengths, advancing in the same direction, and the same explanation has been developed by Lord Rayleigh.

If two infinite trains of waves be represented by $\cos k(vt - x)$ and $\cos k'(v't - x)$ where $k = 2\pi/\lambda$ and $k' = 2\pi/\lambda'$, their resultant is

$$\cos k(vt - x) + \cos k'(v't - x)$$

$$= 2\cos\left(\frac{k'v' - kv}{2}t - \frac{k' - k}{2}x\right)\cos\left(\frac{k'v' + kv}{2}t - \frac{k' + k}{2}x\right)$$

$$= 2\cos\tfrac{1}{2}\alpha\cos[k(vt - x) + \tfrac{1}{2}\alpha] \text{ where } \alpha = d(kv)\,.\,t - xdk.$$

Now, if $k' - k$ and $v' - v$ be very small, we have a train of waves whose amplitude varies slowly from one point to another between the limits 0 and 2, forming a series of groups separated from one another by regions comparatively free from disturbance.

At any time $t$ the position of the middle of that group which was initially at the origin is given by $\alpha = 0$, or

$$(k'v' - kv)t - (k' - k)x = 0.$$

Hence the velocity of the group is

$$\text{V} = \frac{x}{t} = \frac{k'v' - kv}{k' - k},$$

or in the limit when the number of waves in each group is infinite, the relation between the group velocity V and the wave velocity $v$ is

$$\text{V} = \frac{d(kv)}{dk} = \frac{d(1/\text{T})}{d(1/\lambda)} \quad . \quad . \quad . \quad . \quad . \quad . \quad (1)$$

remembering that $kv = 2\pi/\mathrm{T}$ and that $k = 2\pi/\lambda$, or, as it may also be written,

$$\frac{\mathrm{V}}{v} = 1 + \frac{d \log v}{d \log k} = 1 - \frac{d \log v}{d \log \lambda}.$$

Thus if $v \propto \lambda^n$

$$\mathrm{V} = (1 - n)v.$$

Cases—

| | | |
|---|---|---|
| $v \propto \lambda$, | $\mathrm{V} = 0$, | Reynolds's disconnected pendulums |
| $v \propto \lambda^{\frac{1}{2}}$, | $\mathrm{V} = \frac{1}{2}v$, | Deep-water gravity waves |
| $v \propto \lambda^0$, | $\mathrm{V} = v$, | Aërial waves, etc. |
| $v \propto \lambda^{-\frac{1}{2}}$, | $\mathrm{V} = \frac{3}{2}v$, | Capillary water waves |
| $v \propto \lambda^{-1}$, | $\mathrm{V} = 2v$, | Flexural waves. |

The theory of capillary water waves ($\mathrm{V} = \frac{3}{2}v$) has been given by Thomson (*Phil. Mag.* November 1871). Their wave length is so small that the force of restitution due to capillarity largely exceeds that due to gravity. The flexural waves ($\mathrm{V} = 2v$) are those corresponding to the bending of an elastic rod or plate (Rayleigh, *Theory of Sound*, § 191).

Professor Osborne Reynolds (*Nature*, 23rd August 1877 ; also *Brit. Assoc.* Plymouth) gave a dynamical explanation of the fact that a group of deep-water waves advances with only half the rapidity of the individual waves. In this case the energy propagated across any point, when a train of waves is passing, is only one-half of the energy necessary to supply the waves which pass in the same time, so that if the train of waves be limited it is impossible that its front can be propagated with the full velocity of the waves, as this would imply the acquisition of more energy than can in fact be supplied. Professor Reynolds did not contemplate the cases where more energy is propagated than corresponds to the waves passing in the same time, but his argument applied conversely to the results already given shows that such cases must exist. The ratio of the energy propagated to that of the passing waves is $\mathrm{V}/v$. Thus the energy propagated per unit time is $\mathrm{V}/v$ of that existing in a length $v$, or V times that existing in a unit length. Accordingly

$$\frac{\text{Energy propagated per unit time}}{\text{Energy (average) per unit length}} = \frac{d(kv)}{dk} = \frac{d(1/\mathrm{T})}{d(1/\lambda)}.$$

[Lord Rayleigh, Note on "Progressive Waves" (*Proceedings of the London Mathematical Society*, vol. ix. No. 125 ; also *Theory of Sound*.]

## CHAPTER III

### ON THE APPROXIMATE RECTILINEAR PROPAGATION OF LIGHT

**49. The Principle of Huygens—Secondary Waves—Wave Envelopes.**—The first function of any theory of light is to account satisfactorily for the so-called rectilinear propagation of light in a homogeneous medium. The commonest observations on the shadows cast by opaque objects or on the beams of light transmitted through apertures show roughly that light is propagated in right lines or *rays*, and, as we have already mentioned, this apparent rectilinear propagation would seem at first sight to argue strongly in favour of the corpuscular theory. When the facts are more closely scrutinised, however, it is found that the rectilinear propagation is only approximate, for when the source of light is small, such as a narrow slit, then the shadow of an opaque obstacle (a wire, for example) is not so wide as the obstacle, for the light bends round the edges of the obstacle into the geometrical shadow, just as sound bends round corners, only in a much smaller degree.

Hence any proposed theory must account, not for an accurate rectilinear propagation, but for an approximate rectilinear propagation in which there is a slight bending round corners. The apparent superiority of the emission theory consequently breaks down, and in the explanation of this and other phenomena connected with the passage of light through narrow apertures, and past the edges of opaque obstacles, the wave theory obtains a decided advantage.

It is well known that large obstacles cast more or less distinct sound shadows, still in such cases there is considerable bending round corners, but the question at once arises as to how far the possibility of observing this bending depends on the wave length. If the wave length is very small will the amount of observable bending be also very small? It will be seen from the following considerations that this question must be answered in the affirmative, and that the approximate rectilinear propagation of light is a consequence of the minuteness of the wave length.

According to the wave theory each point of a luminous body is a centre of disturbance, or a source from which waves are propagated in the ether. It is therefore necessary to form some conception of the manner in which a wave is propagated.

Let O (Fig. 23) be a centre of disturbance, and let *ab* be the front of a spherical wave diverging from it. The radius of the wave increases with the velocity of light, so that the disturbance now at *ab* will an instant later be at *a'b'*, and the single wave *ab* in travelling out will disturb all the elements of the medium over which it passes. Thus the disturbance of any one element of the medium causes a subsequent disturbance of all the other elements, and we may regard the displacements of the elements of the wave front *ab* as the cause of the subsequent displacement of the elements of *a'b'*. With M, M', etc., as centres, describe a series of equal small spheres[1] to represent the wavelets developed by the centres of disturbance M, M', etc. All these small spheres touch a sphere *a'b'* having its centre at O, and in this manner we get a new wave front *a'b'*. In this manner the wave *ab* is propagated to *a'b'* by the small secondary waves arising from the front of the original wave *ab*. The envelope of these secondary waves is the grand wave, and in this view of wave envelopes Huygens regarded a wave as being propagated. The effective part of each secondary wave in generating the primary wave he supposed confined to that portion of it which touches the envelope.

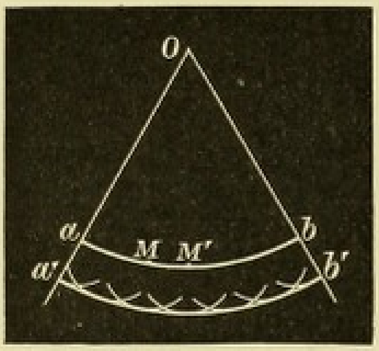


Fig. 23.—Propagation by Secondary Waves.

The energy of *ab* is thus handed on to *a'b'*, and in the same manner from *a'b'* to *a''b''*, etc. Consequently each point of the wave front in any of its anterior positions may be considered as a centre of disturbance from which a secondary wave diverges, and the aggregate effect of these *secondary* waves at any point must be the same as that of the primary wave itself, when passing through that point.

Or, more generally, if round the origin O any ideal closed surface be described, the whole action at any point outside this surface, of the waves diverging from O, may be regarded as due to the motion propagated across the various elements of the surface—that is, each element of the surface, as a centre of disturbance, sends waves to the point under consideration, and the disturbance there is the resultant of their combined action. The general problem, therefore, is to determine the nature of the disturbance at any point, when the full

[1] This assumes the medium to be isotropic.

particulars of the displacement at each point of any surface surrounding the origin O are known. It is convenient, however, to choose the wave front as the surface of resolution, for the vibrations at it at any instant are all in the same phase.

Thus if *ab* be the front of a wave diverging from O, then all points on the surface of *ab* are in the same phase of motion at any instant, and, as time progresses, each point passes through all the phases of the vibration originated by O. Each point of *ab* must, therefore, be regarded as the origin of an *infinite train* of waves, and in calculating the displacement at any point outside *ab*, we must consider all the waves which reach that point *simultaneously* from the various points of *ab*. These waves will have left the various points of *ab* at different times previously, and will, therefore, be in different phases when they reach the point. The general problem, therefore, is to find the resultant of a system of vibrations of different amplitudes and phases, determined by the distances of the various points of *ab* from the point in question.

In order to make this clearer let AB (Fig. 25) represent the trace of a plane parallel to the front of a system of plane waves travelling towards O. At any instant the various points $M_1$, $M_2$, $M_3$, etc., of AB are in the same phase of motion, and this phase changes periodically as the time progresses. Each point of AB, regarded as a centre of disturbance, is the origin of an infinite train of waves (which diverge as spheres around it if the medium be isotropic), so that at any instant the displacement at O is the resultant of an infinite series of disturbances propagated from the various points of AB. The phases of these disturbances are determined by the distances of O from the various points of AB, and their amplitudes depend upon the inclination to the wave front of the line joining O to the corresponding centre of disturbance. Thus, if $v$ be velocity of propagation, the waves reaching O at any instant are—the wave which left P at the time $OP/v$ previously, the wave which left $M_1$ at the time $OM_1/v$ previously, the wave which left $M_2$ at the time $OM_2/v$ previously, etc. etc., so that if the variation of phase with time be known, then the phases of all the disturbances reaching O at any instant are known, and their amplitudes can also be calculated if the law of variation of amplitude be known, so that the resultant effect at O can be estimated.[1]

Interference.

[1] Huygens endeavoured to explain the rectilinear propagation of light by the principle of wave envelopes alone. Thus if *ab* (Fig. 23) be an aperture, and O*ab* a cone of light falling on it from O, he *assumed* that the effect of each element of *ab* in generating *a'b'* was confined to the apex of the secondary wave or that part of it which touches the envelope. This practically assumed the whole question, and it

The question now arises as to what is to limit the secondary waves diverging from each point of the primary wave as origin. In an isotropic medium the disturbance of each little wave spreads generally in a spherical form. Are we to suppose the sphere completed so that each small wavelet is propagated backwards as well as forwards?

We know that a single wave may be transmitted along a stretched cord, or through the air. In this case the agitation of any point causes the future agitation of those in advance of it, but of none in the direction from which it has been propagated. Thus in Fig. 7 the wave travelling along the cord AB disturbs in turn all those parts of the cord in advance of it, but has no effect on those behind it. For this reason it has been supposed by some that each point of the primary wave sends forward only a hemispherical wave, viz. that half of the secondary spherical wave which lies in front of the primary wave. Further, if the disturbance is zero at the base of this hemisphere, it is natural to suppose that the intensity of the disturbance diminishes gradually from the vertex towards the base. This would be the case, for example, if the intensity in any direction varied as the cosine of the obliquity, or inclination to the wave normal.

Stokes's Law.

The law which determines the intensity at each point of a secondary wave has been investigated by Sir G. G. Stokes.[1] He has shown that the effect of an elementary wave at an external point varies as $(1 + \cos\theta)$ where $\theta$ is the obliquity, or angle between the wave normal and the line joining the point to the centre of the elementary wave. This factor will vanish only when $\theta = \pi$, that is for points directly behind the wave. According to this law the secondary wavelet must be regarded as a complete sphere, the disturbance varying gradually from a maximum at its forward apex to zero at the diametrically opposite point in its rear. The effect produced at any point by a wave element depends also on the direction of vibration in the element. Professor Stokes finds it proportional to the sine of the angle between this direction and the radius vector to the point.

It will be sufficient, however, for our present purpose to admit that the effect produced at any point by a secondary wave depends upon the obliquity, and diminishes as the obliquity increases.

was left to Fresnel (*Œuvres*, tom. i. p. 174) to show that the secondary waves diverging from the various points of *ab* must be taken into account, and that these by mutual interference may neutralise each other in some regions of space, and hence produce darkness and shadow.

[1] On the Dynamical Theory of Diffraction, *Math. and Phys. Papers*, vol. ii. p. 243; *Camb. Phil. Trans.* vol. ix. p. 1, 1849.

**50. Definition of Poles and Half-period Elements.**—The *pole* of a wave with respect to any external point O is that point of the wave which is nearest to O, or more accurately, that point of the wave from which the disturbance reaches O in the least time. If the wave is plane the pole of O is the foot of the perpendicular drawn from it to the wave front, and if the wave is spherical, the pole is that point where the wave is intersected by the line joining O to the centre of the sphere.

In general, if the wave be neither plane nor spherical, we define the *pole* (*or poles*) of a wave as the point (or points, or continuous locus of points) on the wave front, from which light is propagated to O in either a minimum or a maximum time.

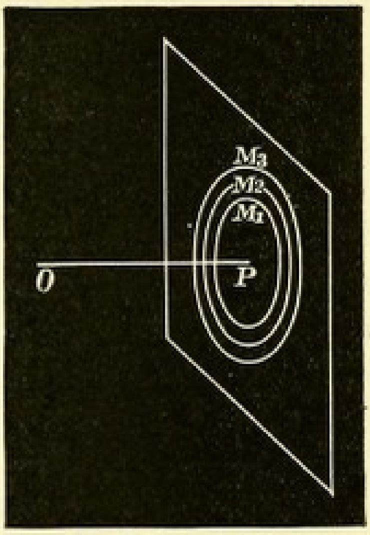


Fig. 24.—Half-period Elements.

Let P (Fig. 24) be the pole of a wave with respect to the point O, and denote OP by $b$. The disturbance from P reaches O sooner than the disturbance from any other point of the wave. With O as centre and a radius $b + \frac{1}{2}\lambda$ describe a curve on the wave front. This curve is the intersection of the wave front with a sphere of radius $b + \frac{1}{2}\lambda$ and centre O, and will include a small element of area on the wave front around P which we shall call the first *half-period element* or *zone*.

With O as centre and a radius $b + \lambda$ describe another sphere meeting the wave front in a second curve which with the first curve will include a little annulus or strip or area on the wave front. This area we term the second half-period element. Similarly by describing other spheres of radii $b + 3\frac{\lambda}{2}$, $b + 4\frac{\lambda}{2}$, etc., we obtain curves on the wave front including the 3rd, 4th, etc., half-period elements, the $n$th half-period element being intercepted between the spheres of radii $b + (n-1)\frac{1}{2}\lambda$ and $b + n\frac{1}{2}\lambda$. A half-period element with respect to any point O is then a narrow strip of the wave front surrounding the pole P of that point, and such that the difference of the distances of its inner and outer edges from O is half a wave length.

**51. Comparison of two Consecutive Half-period Elements.**—Let us now consider the mutual effect of two consecutive half-period elements, $M_1M_2$ and $M_2M_3$, at the point O (Fig. 24). For this purpose let the annulus $M_1M_2$ be divided into a series of concentric circular rings of infinitesimal area. This may be supposed to be done by a system of spheres described round O as centre with radii $r$, $r + \delta$,

$r+2\delta$, $r+3\delta$, etc., where $\delta$ is a very small quantity (a submultiple of $\frac{1}{2}\lambda$) and $r$ is the distance of the inner edge of $M_1M_2$ from O. Then, if $M_2M_3$ be divided up in the same manner, the first ring of $M_1M_2$ will be half a wave length nearer O than the first ring of $M_2M_3$, and being consequently opposite in phase they will neutralise each other at O if the amplitudes of the vibrations which they transmit are equal. In the same way the second ring of $M_1M_2$ will be opposite in phase to the second ring of $M_2M_3$, the third to the third, etc., so that the consecutive strips of one half-period element tend to neutralise the corresponding strips of the other, and if their amplitudes are equal they will destroy each other completely.

Corresponding rings.

We have now to examine how far equality in this respect is realised and on what quantities approximate equality depends. For this purpose it is necessary to remember that the elementary rings into which the half-period elements have been divided are such that they are at uniformly increasing distances $r$, $r+\delta$, $r+2\delta$, etc. from O, and they consequently send vibrations of uniformly increasing phase to O. Rings constructed in this manner will not be rigorously equal in area, and the amplitudes of the vibrations they transmit to O may differ accordingly. In order to find an expression for the area of any such ring let XX′ (Fig. 25) be the width of the annulus, and let $OX' - OX = \delta$ where $\delta$ is a very small quantity. Then if $x$ be the mean radius of the annulus its area is $2\pi x XX'$; but by similar triangles $XX' : \delta :: r : x$, where $r$ is the mean distance of the strip from O, hence $x \, . \, XX' = r\delta$, and the expression for the area becomes

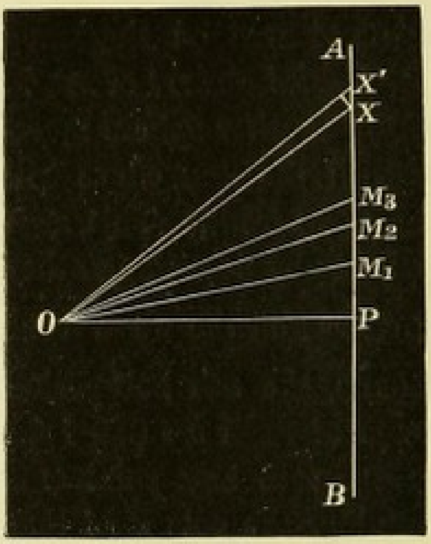


Fig. 25.

$$2\pi r\delta.$$

This expression shows us that if the rings are so constructed that $\delta$ remains constant, then their areas increase as they recede from the centre; but the increase of amplitude arising from this increase of area is exactly counterbalanced by the diminution of amplitude arising from increase of distance from O, if we take as a first assumption that the amplitude of the vibration transmitted to O by any ring is proportional to the area of the ring and inversely as its mean distance from O. Since the area of the annulus is $2\pi r\delta$, this assumption leads to the conclusion that the amplitude of the vibration due to any annulus is proportional to

$$2\pi\delta,$$

and therefore the same for all the rings. The corresponding rings of two consecutive half-period elements would thus produce equal and opposite effects at O and would appear to neutralise each other completely, but as yet we have not introduced any consideration as to how far the amplitude of the vibration due to any element depends on the obliquity—that is, upon the inclination of the wave normal to the line joining O to the element in question. At present we shall merely assume that the effect of obliquity is to diminish the amplitude so that the effect produced at O by any element diminishes as the radius of the ring increases. The result of this assumption is that the elementary rings of the half-period element $M_1M_2$ are more effective at O than the corresponding rings of $M_2M_3$, and instead of complete neutralisation we have an outstanding difference for every consecutive pair of zones, and the resultant of these outstanding differences is the integral effect of the wave.

Effect of obliquity.

It is interesting to notice the effect of this diminution of amplitude on the resultant of a single half-period element. If such an element be divided, as above, into a great number of concentric rings of equicrescent phase, and if there were no diminution of amplitude arising from increase of obliquity, then, when represented graphically, the resultant of the whole zone would be represented by the diameter OP of a circle OMP, as shown in Fig. 15, p. 50, the phase of the resultant being 90° behind that of the vibration coming from the inner edge of the zone. The effect of diminution of amplitude, however, is such that in constructing the curve OMP the elements of length diminish as we proceed along it by revolving the tangent through equal increments of angle. This leads to an increase in the curvature of the amplitude curve corresponding to any given portion of the zone, and, as a consequence, the phase XOP (Fig. 14) of the resultant is less than 90° behind that of the inner edge of the zone. The phase of the resultant of a single half-period element may thus be less than that of the central element of the zone and may correspond to a ring of the zone situated between the central ring and the inner edge.

Before concluding the comparison of two consecutive half-period elements it will be useful to examine the outstanding difference arising from obliquity in so far as it depends on the wave length. We have found that the amplitude of the vibration due to any elementary ring is proportional to $2\pi\delta$, when the obliquity is neglected, and if we denote the obliquity by $\theta$ and the distance OP by $p$ we have $\cos\theta = p/r$, and consequently any function of $\theta$ may be expressed in terms of $r$ and the constant $p$. Hence, if the law of variation of amplitude with obliquity be known, it may be expressed as a function of $r$, and the amplitude of the vibration due to any elementary ring may be written in the form

$$A = af(r).$$

The amplitude of the vibration due to the corresponding ring of the consecutive zone will therefore be

$$A' = af(r + \tfrac{1}{2}\lambda) = af(r) + \tfrac{1}{2}a\lambda f'(r) + \text{etc.}$$

As these are opposite in phase their joint effect will be

$$A' - A = \tfrac{1}{2}a\lambda f'(r) + \text{etc.},$$

Any law. so that if $\lambda$ is small the outstanding difference will be a small quantity proportional to $\lambda$, viz.—

$$\tfrac{1}{2}a\lambda f'(r).$$

Hence when $\lambda$ is small, the resultant of two consecutive zones is small compared with either.

For example, if we assume that the amplitude varies simply as the cosine of the obliquity, we have $f(r) = p/r$ and $f'(r) = -p/r^2$, so that

Cosine law.

$$A - A' = \tfrac{1}{2}ap\lambda/r^2.$$

Thus with this law of variation (or with Stokes's law, p. 62) the outstanding difference varies directly as $\lambda$ and inversely as the square of the distance; in other words, as the zones recede from P their effects become smaller and more nearly equal to each other.

**52. Plane Wave.**—In order to estimate the whole effect of a plane wave at any point O we divide it, as in the foregoing article, into a system of half-period elements with respect to O. Let $m_1$ denote the resultant effect at O of the first half-period element, $m_2$ that of the second, and so on. Then since the corresponding rings of two consecutive half-period elements are opposite in phase at O, the whole effect of the plane wave is represented by the series

$$S = m_1 - m_2 + m_3 - m_4 + \text{etc.}$$

The only definite knowledge which we have as yet deduced concerning this series is that the consecutive terms are very nearly equal and gradually grow less and less in absolute value as they recede from the beginning of the series, and setting out with this information, it may be easily shown that its sum S approximates in the limit to the value $\tfrac{1}{2}m_1$. For since the successive terms decrease continuously according to some law, from the beginning to the end of the series, it follows that if ordinates $Om_1$, etc. (Fig. 26), be erected at equal distances along a right line OX, so that the lengths of the ordinates represent the terms $m_1$, $m_2$, $m_3$, etc., of the series S, then a smooth curve may be drawn through the extremities of these ordinates which will be either concave or convex towards the axis OX, according to the law of variation of the constituents of the series. Now obviously $m_1 a = m_1 - m_2$, and $m_3 b = m_3 - m_4$, and consequently S is the sum of the intercepts $m_1 a$, $m_3 b$, etc., and this sum includes the alternate steps only of the stairs leading down from $m_1$ to X. But the sum of all the steps $(m_1 - m_2) + (m_2 - m_3) +$ etc. is obviously $m_1$, hence the sum of the alternate steps must in the limit be half of the whole—that is

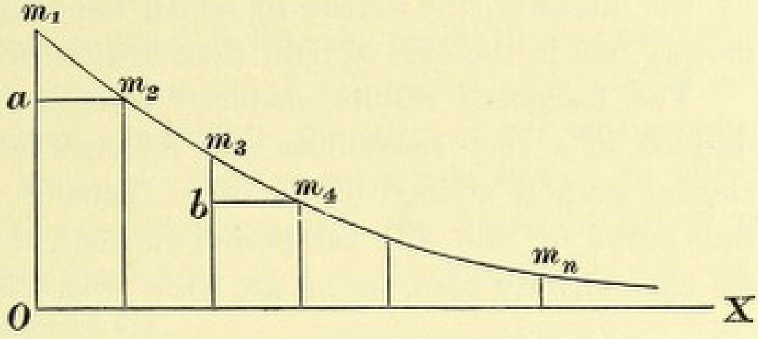



Fig. 26.

$$S = \tfrac{1}{2} m_1.$$

In deducing this result the only condition that has been introduced is that the terms gradually diminish to zero, so that the curve gradually approaches, and ultimately touches the axis OX.

The same reasoning proves that the sum of all the terms after the $n^{th}$ approximates to $\frac{1}{2} m_n$, and consequently the error introduced by neglecting all the terms after the $n^{th}$ does not exceed $\frac{1}{2} m_n$. Maximum error.

Now, if the wave length be very small, a small area around P will contain a large number of half-period elements, and the resultant effect of this portion of the wave will be approximately the same as that of the complete wave, the error being less than half the value of the last half-period element, which, when $n$ is large, becomes vanishingly small. We conclude, therefore, that when $\lambda$ is small, the *effective* portion of the wave is confined to a small area around the pole of the wave. If this area be intercepted by an opaque obstacle the remainder of the wave will have no appreciable effect at O—that is, a small obstacle at P will screen O almost entirely from the wave, and this is what we mean when we say that light is propagated in right lines.

The approximate rectilinear propagation of light is, therefore, a consequence of the extreme shortness of the wave length, and is explained by the principle of interference combined with Huygens's supposition concerning secondary waves. When the wave length is large, as in the case of sound, the bending round corners becomes very noticeable, but in this case also fairly distinct shadows are cast, and screening occurs, when the obstacle is large compared with the wave Bending.

length. The only difference is that when the wave length is small the intensity falls off much more rapidly as we recede within the geometrical shadow, and as the limits of observation are determined by the intensity, the extent to which bending is observable is enormously less in the case of light than in that of sound.

Graphic method.

The foregoing results, deduced from the series S, may be also derived in a very convenient and instructive manner by aid of the elegant graphic method of Art. 45. Thus, if we construct the amplitude curve for the first half-period element, it will be represented, as we have already seen, by an arc $Oa'a$ (Fig. 27), which is very nearly

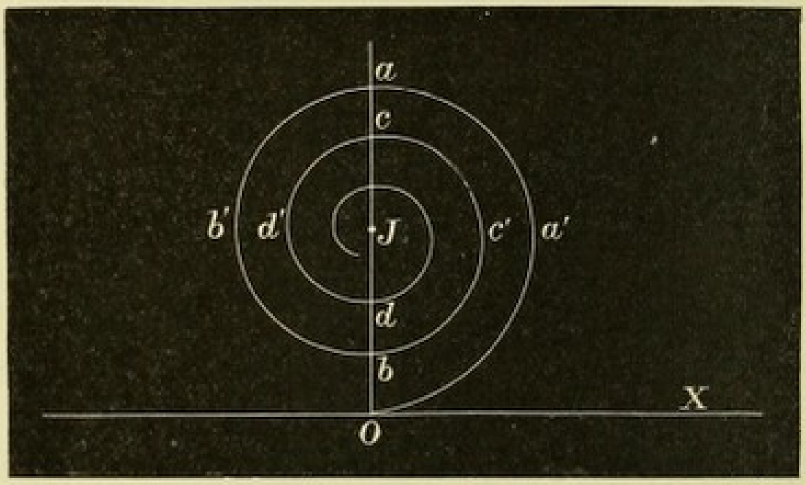



Fig. 27.—Vibration Spiral.

a semicircle, and in the same way if the construction be extended the second half-period element will be represented by an arc $ab'b$, which is also nearly a semicircle. The resultant of the first half-period is

$$m_1 = Oa,$$

and the resultant of the second is

$$m_2 = ab,$$

while the resultant effect of the pair taken together is

$$m_1 - m_2 = Ob,$$

the difference $Ob$ being small compared with either $m_1$ or $m_2$. Similarly, if the construction be extended to the other half-period elements, so as to embrace the whole wave front, the third half-period will give rise to the arc $bc'c$, the fourth to $cd'd$, etc. etc. Hence the complete curve representing the whole wave is a spiral of an infinite number of nearly circular convolutions of ever-decreasing radius surrounding a point J on the line $Oa$, very approximately halfway between O and $a$.

Since OJ represents the amplitude of the vibration excited by the whole wave, and since $Oa = m_1$, it follows that in the limit we have

$$S = \tfrac{1}{2}m_1.$$

So also the figure informs us (as does also the series S) that the effect of the first two, or any even number of half-period elements, is less than the effect of the whole wave, while the effect of any odd number of elements is greater than the effect of the whole wave. In either case, as the number of elements is increased, the deficit or excess of effect over OJ gradually diminishes—that is, the lengths $Oc$, $Od$, etc., become more nearly equal to OJ.

We are consequently led to the conclusion that if the whole wave is screened off, except the first half-period element, the intensity at the point O under consideration will be about four times as great as that produced by the whole wave, whereas, if the aperture be increased so as to transmit two half-period elements, the intensity at O will be reduced almost to zero. By increasing the aperture so as to transmit three half-periods the intensity again rises to a maximum, and by further increasing it so as to transmit four elements the intensity falls to a minimum. By still further increasing the aperture the intensity at O passes through a succession of maxima and minima, but these become less and less pronounced as the number of elements is increased, so that after the aperture reaches a certain limit further increase will produce no noticeable effect in the illumination at O. Now a large number of these elements is included in a small area around P, the pole of O, and we consequently conclude that the effective portion of the wave is restricted to a small area around P in so far as the illumination at O is concerned.

Effect of screening.

On the other hand, if we consider the effect of placing a small screen, instead of an aperture, at the pole of the wave, we see at once that when the screen just covers the first zone, the amplitude of the vibration at O will be represented by $a$J, for the part $Oa'a$ of the amplitude curve is cut off while the remainder of the spiral remains effective. Similarly, if the screen covers two elements, the intensity at O will be represented by the square of $b$J, and so on. We conclude, therefore, that as the screen increases in size the intensity at O gradually diminishes to zero without passing through maxima or minima, and that when the screen is large enough to cover a considerable number of zones the illumination at O falls below the limits of observation.

These theoretical deductions are fully confirmed by the results obtained by experiment, as will be seen later on, and it may be well to remark that there is no complementary relation, such as might at first sight be imagined, connecting the intensity at O when part of the wave is transmitted through an aperture with that obtained when the same part is obstructed by a screen, and the remainder transmitted; for, in the former case, the illumination passes through maxima and minima as the size of the aperture is varied, whereas, in the latter case, no such maxima or minima occur.

**53. Spherical Wave.**—In the same manner we may calculate the effect of a spherical wave at an external point. Let C (Fig. 28) be the centre of the wave, P the pole of O, and let OM′ − OM = $\delta$, where $\delta$ is a very small quantity.

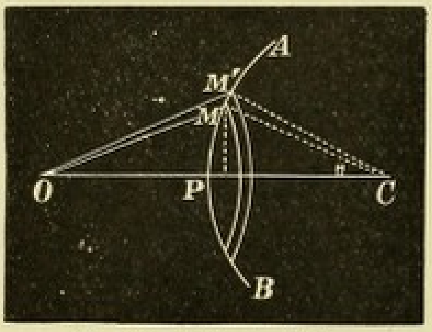


Fig. 28.—Spherical Wave.

The area of the annulus intercepted on the surface of the wave by two spheres described round O with radii OM and OM′ is

$$2\pi a \sin\theta \,.\, \mathrm{MM'},$$

where $\theta$ denotes the angle OCM. For since MM′ is very small the annulus is a circle of mean radius $a \sin\theta$, viz. the perpendicular from M on OC, $a$ being the radius CM of the wave.

Now if N be the foot of the perpendicular from M on OM′,

$$\frac{\mathrm{NM'}}{\mathrm{MM'}} = \sin \mathrm{OM'C}, \therefore \mathrm{OM'}\frac{\mathrm{NM'}}{\mathrm{MM'}} = \mathrm{OM'} \sin \mathrm{OM'C} = \mathrm{OC} \sin \mathrm{OCM'},$$

or since the angle MCM′ is very small and NM′ = $\delta$, we have

$$r\delta = (a+b)\mathrm{MM'} \sin\theta,$$ [1]

consequently the area of the annulus is

$$2\pi r\delta \frac{a}{a+b}.$$

For a plane wave $a = a + b = \infty$, and the area becomes as before $2\pi r\delta$.

Hence if we neglect the variation of amplitude arising from obliquity and take it to vary directly as the area of the strip and inversely as the distance $r$, we find that the amplitude produced at O by any annulus, such that the difference of the distances of its inner and outer edges from O is $\delta$, is simply proportional to [2]

$$2\pi\delta \frac{a}{a+b},$$

[1] Or thus: $r^2 = a^2 + (a+b)^2 - 2a(a+b)\cos\theta$, $\therefore$ $rdr = a(a+b)\sin\theta d\theta$, but here $dr = \delta$ and $\mathrm{MM'} = ad\theta$, therefore at once $r\delta = (a+b)\,\mathrm{MM'}\sin\theta$.

[2] Or if $\phi$ is the phase difference, the amplitude is proportional to $\lambda\phi a/(a+b)$. For a given phase difference the intensity is proportional to $\lambda^2$.

and consequently all rings for which $\delta$ is the same produce effects of equal magnitude at O. We find, therefore, as before that when the effect of obliquity is neglected the consecutive half-period elements destroy each other at O, but when the obliquity is taken into account the effects of the various zones at O gradually diminish as they recede from the pole P, and the whole effect is the sum of the series

$$S = m_1 - m_2 + m_3 - m_4 + \text{etc.},$$

in which the terms gradually diminish from left to right. From this stage all the reasoning of the foregoing article applies together with graphic construction, and we conclude that the whole effect is approximately the same as half that of the first half-period element, or

$$S = \tfrac{1}{2}m_1.$$

**54. Wave of any Form.**—When plane or spherical waves are reflected or refracted at curved surfaces the wave front in general becomes of a more complicated character; hence, in order to complete the problem of the rectilinear propagation of light in isotropic media, it is necessary to consider the case in which the wave front is a surface of any form. The first point to be remarked about such a surface is that it may present several poles with respect to any point O, for it may be such that the radius vector drawn from O to a variable point on the surface passes through several maxima and minima as the point traverses the surface. These poles are the points in which the surface is touched by a sphere of variable radius described round O, and they may be isolated from each other, or they may in some cases be consecutive points on the surface and form a continuous locus or curve of poles. Pole locus.

Now since the radius vector from O to any pole P (Fig. 29) is either a maximum or a minimum, it follows that the line OP is a normal to the surface at P. For this reason the element of surface in the immediate neighbourhood of P will have sensibly the same effect at O as the corresponding element around P taken either on the tangent plane or on the sphere of closest contact with the surface at P. Hence when the wave length is very small, so that a considerable number of

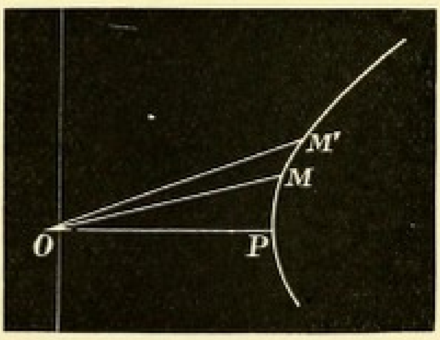


Fig. 29.—Wave of any Form.

half-period elements are contained in a small area around P, this portion of the wave will produce an effect at O which will not be sensibly increased by increasing the magnitude of the area in question, or if this portion of the surface be intercepted by a screen, the darkness at O will not be sensibly increased by increasing this size of the screen.

As we recede from P the obliquity increases, and as before we can see from general considerations that two consecutive half-period elements at a distance from P very approximately neutralise each other. For if we consider two consecutive half-period elements intercepted on the surface by spheres described round O with radii $r - \frac{1}{2}\lambda$, $r$, and $r + \frac{1}{2}\lambda$, then, if $\lambda$ be small, any two neighbouring portions of these elements will be related to each other as the neighbouring portions of two consecutive half-period elements on a plane or spherical surface, at a distance from the pole such that the obliquity is the same. When this obliquity is sensible the outstanding difference of effect between two consecutive half-period elements will consequently be vanishingly small when $\lambda$ is small, and the whole effective portion of the wave will be limited to a small area surrounding each pole.

In the general case, therefore, the whole illumination at O will appear to come from the immediate neighbourhood of certain points on the surface, and this illumination will not be appreciably influenced by screening off the remainder of the surface. These points may be isolated and can be treated as separate sources when they are removed from each other by a large number of half-period elements, the effect of each being approximately the same as half that of the first half-period element, or they may be close together and form continuous loci on the surface, so that the illumination may appear to come to O from certain curves traced on the surface.

# CHAPTER IV

## REFLECTION

**55. Reflection, Regular and Irregular.**—When light falls upon the surface of separation of two media, part of it is generally turned back or *reflected*. Thus when a pencil of light, admitted into a darkened room by a hole in the shutter, is allowed to fall upon a polished metallic mirror, a reflected beam may be seen[1] leaving the mirror and travelling along a certain definite path. This portion of the light is said to be *regularly* reflected, in contradistinction to another portion of the light, which, after falling upon the mirror, is scattered at the surface in all directions—or *irregularly* reflected. This scattering is due to the inequalities of the reflecting surface, and it diminishes as the polish of the surface is made more perfect.[2]

It is by means of this scattered light that we see most of the bodies around us which are not self-luminous. Thus if the light were all regularly reflected from a mirror the eye would be affected only when placed in the reflected beam, and then a bright image of the sun would be seen in the mirror. If the eye were elsewhere, no light would enter it and nothing would be seen. The scattered light, however, is diffused in all directions from ordinary objects and enters the eye from all parts of the surface, so that they can be seen in every position of the eye.

In speaking of reflected light in future we mean that light which is regularly reflected according to the laws which we are about to enunciate. The so-called irregularly reflected light, when there is any

[1] By means of reflecting dust particles in the air which scatter the light in all directions. We cannot, of course, see light travelling through space. It only affects the eye when it enters it, and therefore must either enter it directly from the source or be reflected into it.

[2] The expression "irregularly reflected" is here rather an abuse of terms. There is no irregularity in the reflection. The irregularity is confined to the reflecting surface. The ordinary laws of reflection are obeyed in full, and are not departed from unless the linear dimensions of the reflecting surface, or of the rugosities on it, are small compared with the wave length.

occasion to speak of it, will be specially referred to as scattered or diffused light.

**56. Laws of Reflection.**—The beam of light falling upon the mirror is termed the *incident light*, and the angle which its direction makes with the perpendicular (or, as it is often called, the *normal*) to the surface at the point of incidence is named the *angle of incidence*, while that part of the light which is reflected is known as the *reflected light*, and the angle which its direction makes with the normal to the surface is the *angle of reflection*. The relations between the angles of incidence and reflection have been known from the earliest times, and are stated in the *Laws of Reflection:* "The angles of incidence and reflection are in the same plane, and are equal."

The first part of this statement affirms that the reflected ray lies in the plane containing the incident ray and the normal to the surface at the point of incidence, while the second expresses the equality of the angles of incidence and reflection.

The intensity of the reflected light generally increases with the angle of incidence and with the polish of the surface. It also depends largely on the nature of the medium from which it is incident, and on that from which it is reflected. For example, much more light is reflected, under the same circumstances, from a plate of glass in air than from the same plate immersed in water. The variation of the reflecting power of a surface with the angle of incidence is well illustrated by comparing water and mercury. At perpendicular incidence water reflects about the fiftieth part of the incident light, while mercury reflects about the two-thirds; but at an incidence of $89\frac{1}{2}°$ they each reflect about 72 per cent of the incident light.

An accurate experimental proof of the laws of reflection is furnished by observations with such an instrument as the Meridian Circle. Adjust the telescope to observe a star directly, and then observe the reflection of the same star seen in the horizontal surface of a basin of mercury. The telescope in these two observations will be found to make the same angle, on opposite sides, with the vertical line.[1]

The graduation of such an instrument is the most perfect that can be accomplished by human skill, and yet the smallest divergence from the preceding law has never been detected.

**57. Illustration of Reflected and Refracted Waves.**—If a perfectly elastic ball impinges directly on another of equal mass at rest, the second ball will exactly take up the motion of the first, while the first

[1] [Systematic observations of this kind are made in several observatories. The minute discordance observed has never been attributed to inaccuracy in the law. See the annual volume of *Greenwich Observations*.]

comes to rest at the spot where it impinged on the other. The whole process is as if the first ball moved on through the other without disturbing it or being disturbed itself.

So again, if a number of similar balls be placed in a row (Fig. 30, $\alpha$), and if the one at the end (A) be struck in the direction of the row, it will move forward and impinge on the second and come to rest there, while the second moves forward to strike the third, and comes to rest in turn. In this manner the blow is communicated by each ball to its successor, and the disturbance travels along the whole line, leaving all the balls at rest except the last (B), which moves forward with the velocity and energy initially communicated to the first (A). This is the case of a compressional wave travelling in air (as in sound) along a uniform straight tube. Now let us suppose that after the row AB we have another row A′B′ (Fig. 30, $\beta$) of heavier balls. When the ball B moves forward it strikes A′ and rebounds (since its mass is less than that of A′). At the same time A′ moves forward and strikes its neighbour, which in turn performs its part, and the disturbance travels along the row of larger balls A′B′, each coming to rest when it impinges on its successor, because they are all of the same mass. But the state of things has now altered in the row AB, for B has rebounded from A′, and, travelling backwards, has struck its neighbour and come to rest. This disturbance is handed on from ball to ball as a disturbance along the row in the backward direction BA. Hence the disturbance in the first row (AB) has given rise to two other disturbances,—a direct one in the second set A′B′, and a reflected one in the backward direction BA in the first set.

(α)
(β)

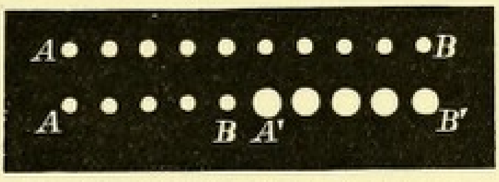


Fig. 30.—Direct and Reflected Waves.

In the same manner we may suppose the disturbance to arise in the row of larger balls. Thus if B′ be struck the impulse will be communicated along the line to A′, and A′ in its turn will move forward and impinge on B, but as B is of less mass than A′, the ball A′ will not come to rest, but will follow after B, and, therefore (if the balls be imagined connected with weightless threads), A′ will pull its successor after it, which in turn will act upon its neighbour, and so on. A second disturbance is set up in the row A′B′, which consists of a further motion of the balls in the same direction as the original disturbance.

We may now liken the two rows of balls to two media separated by a common surface. The smaller lighter balls will correspond to the rarer medium and the heavier balls to the denser. When any

Reflected and refracted waves.

disturbance originates in one of the media it is propagated through it, and when it arrives at the surface of separation two new disturbances are set up, one (refracted) in the second medium, and another (reflected) in the first medium. It therefore follows from the wave theory that when light, travelling in one medium, comes to the surface of another, part of it should penetrate into the second medium, or we should have a refracted wave, and part of it should be propagated backwards in the first medium in the form of a reflected wave.

Since the phenomena of the reflection and refraction of light exist, we must admit that the ether is modified in some way by the presence of matter, and differently in different substances. For example, the ether in glass cannot be in the same condition for vibration as the ether in air or water. Its so-called elasticity and density—that is to say, those properties which enable it to propagate wave motion—are modified by the substance which it permeates.

**58. Deduction of the Two Laws of Reflection.**—Let AA′ (Fig. 31) be the surface of separation of two media, and AB the front of a plane wave incident on it. Each successive portion of the surface as soon as the wave reaches it becomes the centre of two diverging waves, one (reflected) in the upper medium, and the other (refracted) in the lower. These waves travel with different velocities, but if the medium be homogeneous and isotropic the secondary waves will be spherical. At present we shall confine our attention to the reflected wave.

If BA′ and MP be perpendicular to the wave front AB, then B is the pole of A′ and M is the pole of P, so that A′ is illuminated by the element of the wave at B, and P by the element at M; or, in other words, BA′ and MP are in the direction of what we call the rays of light. When the light from B reaches A′ the light from M has arrived at P some time before, and would have reached N if it had not been obstructed by the surface, but on reaching P a reflected wave is developed which diverges into a sphere of radius PM′ = PN, and similarly the reflected wave at A has diverged into a sphere of radius AB′ = AD.

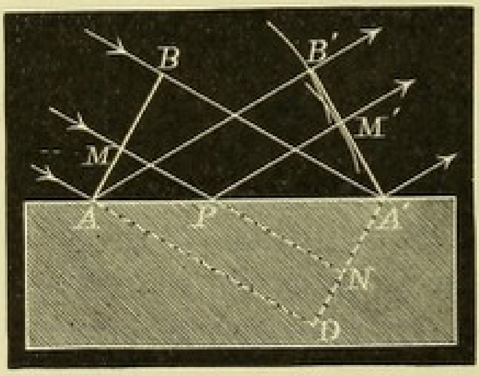


Fig. 31.—Reflection of a Plane Wave at a Plane Surface.

If the plane of the wave AB and the surface of separation AA′ be perpendicular to the plane of the paper, the line through A′ perpendicular to the plane of the paper is the intersection of the surface of separation with the wave front at A′. Through this line draw a plane to touch the reflected wave diverging from A and let the point of contact be B′. Now since AB′ is the radius of this wave, at the instant the light from B reaches A′, it follows that AB′ = BA′ = AD, since the reflected light travels with the same velocity as the incident. Hence the triangle AA′B′ is equal in all respects to the triangle AA′B or to AA′D. Consequently if from P we let fall a perpendicular PM′ on A′B′ we will have PM′ = PN, and therefore the reflected wave diverging from P will touch at M′ the tangent plane A′B′ to the wave from A. Similarly the waves diverging from every point of the surface will touch the same plane. This plane is therefore the reflected wave envelope, and A′B′ is the trace of the reflected wave at the instant the light from B reaches the point A′.

The angle A′AB is the angle between the plane of the incident wave front and the surface; it is therefore equal to the angle between the normal to the wave front, or the ray, and the normal to the surface. Hence A′AB is equal to the angle of incidence. Similarly AA′B′ is equal to the angle of reflection, but these angles are equal by the equality of the triangles ABA′ and AB′A′. The lines AB′, PM′, etc., are the normals to the reflected wave front, that is the reflected rays. Any one of these rays obviously lies in the plane containing the corresponding incident ray and the normal to the surface. The two laws of reflection are thus completely accounted for by the wave theory.

Let us now investigate the matter a little more closely. It appears from what has been already said that the effective portion of the wave AB in illuminating P is confined to a very small element around M, the foot of the perpendicular from P on the wave front. So in like manner, if A′B′ were the incident wave, AB would be the reflected wave, the path being exactly retraced, and P would be illuminated by the element of A′B′ around M′. We should therefore expect that the element M′ of the reflected wave is illuminated by the point P of the surface, or further back still, by the element M of the incident wave. It is easy to show that this is the case.

On the line AA′ find points P, $P_1$, $P_2$, etc. (Fig. 32), such that the path $MP_1M'$ exceeds the path MPM′ by half a wave length, the path $MP_2M'$ exceeds MPM′ by two half wave lengths, etc.; that is, each path exceeds its predecessor by half a wave length. It is well known, of course, that the path MPM′ is less than any other if the lines MP and PM′ are equally inclined to the surface. Therefore

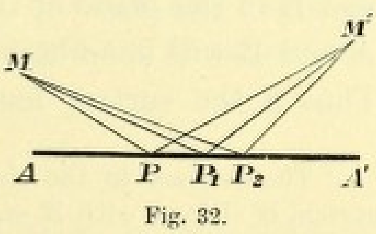


Fig. 32.

the path MPM′ is that along which it takes the light the least time to reach M′ from M after reflection from the surface AB. Therefore in estimating the illumination at M′ by the reflection from AA′ we may consider each point of AA′ as the origin of a disturbance propagated to M′, and find the resultant effect. The surface being divided up into half-period elements, as indicated above, we can easily show as before that of these elements those immediately around P are the greatest, and that the elements diminish[1] rapidly at first and then more slowly till they become practically equal, and being opposed in effect at M′ they produce no illumination there. Consequently the effective portion of the surface AA′ in illuminating M′ is confined to a very small element of the surface at P. If M is a single luminous point an eye placed at M′ will perceive a bright point in the direction M′P. The illumination which reaches M′ from M is propagated in the same time and as if it came from a point situated at an equal distance on the other side of the surface. Similarly every point of the reflected wave is illuminated by a corresponding point of the incident wave as if the light came from the corresponding point of a line through A parallel to A′B′ (Fig. 31). This line is the reflection of AB in the surface. We see then that each point of the reflected wave is illuminated by that point of the incident wave which sends light to it in the least time. This is an example of "the principle of least time," which is of very wide application and fertility in the theory of light.

We have now proved that the disturbance at any point M of the incident wave is propagated along MP, and after reflection at P it travels to M′, a corresponding point on the plane A′B′. The plane A′B′ is the locus of the points which are simultaneously disturbed. It is the reflected wave, while PM and PM′ are rays obeying the laws of reflection enunciated above (Art. 56).

*Cor.*—The time taken by light to travel from any point of the incident wave to the corresponding point M′ of the reflected wave is the same for all rays and a minimum. For $MP + PM' = MN = BA'$.

**59. Reflection of a Spherical Wave at a Plane Surface.**—Let us now consider the case of spherical waves diverging from a centre O and falling upon a plane reflecting surface AB (Fig. 33). Let the wave front at the instant under consideration meet the surface at A and B in the plane of the figure. Each point of the surface between A and B will have become by this time the centre of a reflected wave. Thus if the surface had not been present the wave would have pro-

[1] The decrease in the elements $PP_1$, $P_1P_2$, etc., is well exhibited by describing a system of ellipses with M and M′ for foci and major diameters equal to $l+\frac{1}{2}\lambda$, $l+\lambda$, $l+\frac{3}{2}\lambda$, etc., where $l = MP + M'P$. If the line AA′ is a tangent to the inner ellipse the intercepts made on it by the other conics are the half-period elements.

ceeded unobstructed, and would occupy the position ANB. The effect of the surface, however, is such that M, the foot of the perpendicular from O on AB, has become the centre of a spherical reflected wave of radius MN′ = MN, and any other point P is the centre of a spherical reflected wave of radius PQ′ = PQ.

It is clear that all these reflected waves will touch a sphere of centre O′ and radius O′N′, where MO′ = MO. For join O′ to P, and produce the joining line to meet this sphere at Q′. Then O′Q′ = O′N′ = ON = OQ, and OP = O′P, therefore PQ = PQ′, or PQ′ is equal to the radius of the wavelet diverging from P, and it is also normal to the sphere AN′B. Hence the reflected wave diverging from P touches the sphere AN′B, and this sphere is therefore the envelope of the reflected waves, or the limit to which the reflected disturbance has been propagated when the incident wave meets the surface at A and B. The reflected wave front is consequently a sphere diverging from O′ as centre, or the reflected light appears to diverge from a point O′ on the other side of the surface, and at the same distance from it as O. This point is termed the *image* of O in the surface.

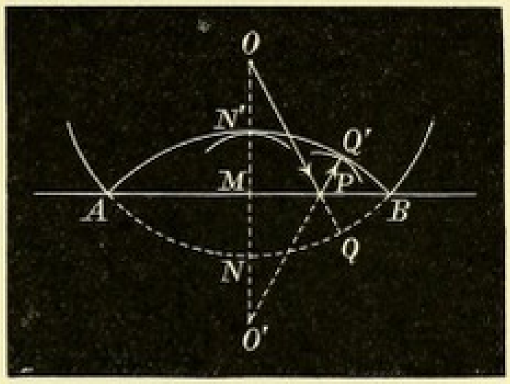


Fig. 33.

Before dismissing this case it may be noticed that the effect of a plane surface in reflecting a spherical wave is simply to reverse its curvature. Thus the incident wave ANB diverging from O is converted into a wave of equal radius diverging from O′.

**60. Measurement of Curvature.**—As the effect of reflection, or refraction, is in general to change the curvature of a wave, it will be convenient to define the measurement of curvature before dealing with waves reflected or refracted at curved surfaces. In the case of a *uniformly* bent curve (such as a circle) the curvature is measured by the bend per unit length—that is, by the angle between the tangents at the extremities of a unit length of the curve. In the case of a circle this is numerically equal to the angle subtended at the centre by a unit length of the circumference, or the same as the ratio of any angle at the centre to the arc subtending it. Thus the general measurement of the curvature of a circle is—

$$\text{Curvature} = \frac{\text{angle}}{\text{arc}} = \frac{1}{\rho},$$

where $\rho$ is the radius of the circle, and the angle is expressed in circular measure.

When the curve is not a circle the rate of bending varies from point to point, but the curvature at any point is still measured by the limit of the above ratio when the arc and angle are taken very small. The curvature of a curve at any point is thus the reciprocal of a length—namely, the radius of curvature, that is, the radius of the circle osculating the curve at the point in question.

The[1] sagitta.

The curvature of a small arc, AB (Fig. 34), may be conveniently[1] expressed in terms of the *sagitta* PM. Thus we have $PM \times MQ = MA^2$, but when PM is small $MQ = 2\rho$ approx., and consequently $PM = MA^2/2\rho$—that is, for an arc of given chord, the sagitta PM is directly proportional to the curvature. And this is what would have been expected, for in the limit PM clearly measures the bulge or bend of the arc. It follows, therefore, that if two arcs, APB and

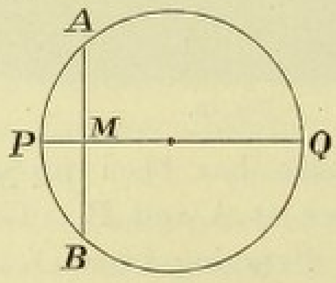


Fig. 34.

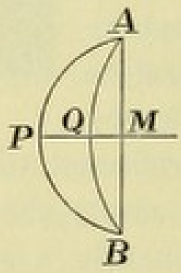


Fig. 35.

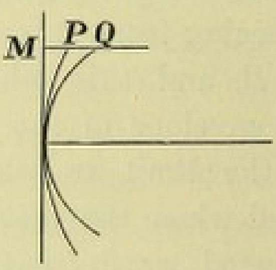


Fig. 36.

AQB (Fig. 35), have the same chord AB, their curvatures are directly as their sagittæ PM and QM. Or if the arcs touch each other, as in Fig. 36, and if a tangent be drawn at the point of contact, then the intercepts PM and QM made by the arcs on a perpendicular to the tangent are proportional to the curvatures of the arcs. For in the limit these intercepts clearly measure the amounts of bend of the arcs.

*Ex.*—If the curvature of a circle be $\sigma$ when its radius is $\rho$, prove that when the radius becomes $\rho + c$ the curvature becomes

$$\sigma' = \frac{\sigma}{1 + c\sigma}.$$

[In this case we have $\sigma = 1/\rho$ and

$$\sigma' = \frac{1}{\rho + c} = \frac{1/\rho}{1 + c/\rho} = \frac{\sigma}{1 + c\sigma}.\Big]$$

**61. Reflection of a Plane Wave at a Spherical Surface.**—We shall now consider the reflection of a plane wave at a spherical surface. Let AMB (Fig. 37) represent the reflecting sphere, and XABY the trace of a plane parallel to the front of the incident wave. As the

[1] This method of treating the problems of reflection and refraction was given in full by Professor S. P. Thompson in 1889 (*Phil. Mag.* vol. xxviii. p. 232), and was partially employed in the first edition of this work.

wave approaches the surface (from left to right) it comes into the position of a tangent plane to the surface, first touching it at some point M. At this instant M becomes the centre of a reflected wave, and as the original wave moves farther to the right, each point of the surface in turn becomes the centre of a reflected wave. Thus when the incident wave occupies the position XY, the part between A and B will have been reflected by the surface into the wave AN′B, such that MN′ = MN, for MN′ is the distance to which the reflected disturbance travels, while the incident wave travels over the distance MN.

Now when the arc AB is small, MN is proportional to the curva-

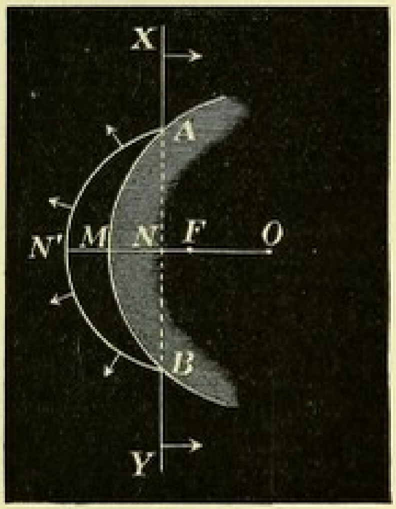



Fig. 37.

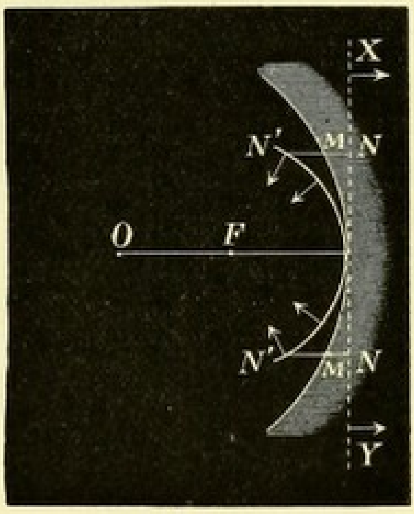



Fig. 38.

ture of the surface and NN′ is proportional to the curvature of the reflected wave. But

$$NN' = 2NM,$$

and we conclude that the curvature of the reflected wave is twice that of the reflecting surface. The action of the surface in reflecting a plane wave is therefore to imprint on the reflected wave a curvature equal to twice the curvature of the surface. In other words, when a plane wave is reflected at a convex spherical surface, the reflected wave diverges from a point F halfway between the centre of the mirror and its surface. This point is called the *principal focus* of the mirror, and the distance MF is termed its *focal length*. The focal length $f$ of a spherical mirror of radius $\rho$ is consequently determined by the equation

Impressed curvature.

$$f = \tfrac{1}{2}\rho.$$

Since $1/f$ measures the curvature impressed on a plane wave by reflection at a spherical surface, this quantity measures the curvature producing power of the mirror and is termed the *focal power*. It is

Focal power.

clear, therefore, that the focal power is equal to twice the curvature of the mirror.

The case of a concave spherical mirror is shown in Fig. 38, in which MM represents the reflecting surface and XY is the trace of a plane parallel to the face of the incident wave, supposed travelling from left to right. If the reflecting surface had not been present XY would represent the incident wave in one of its positions, but by the reflecting action of the surface NN is converted into N′N′, where obviously in the limit MN = MN′, and therefore the curvature of the reflected wave N′N′ is twice that of the reflecting surface. The centre F of N′N′ is consequently halfway between O and the surface, or the action of a concave reflecting surface is to convert a plane wave into a spherical wave, of twice the curvature of the surface, which converges to a point halfway between the centre of the mirror and its surface.

**62. Reflection of a Spherical Wave at a Spherical Surface.**— We have seen that a plane surface in reflecting a spherical wave simply reverses the curvature of the wave, and that a spherical surface imprints twice its own curvature on a plane wave in reflecting it, we might therefore suspect that a spherical surface in reflecting a spherical wave would reverse the curvature of the incident wave, and in addition impress it with twice the curvature of the reflecting surface. That this is the case is very easily proved. Thus if we suppose a

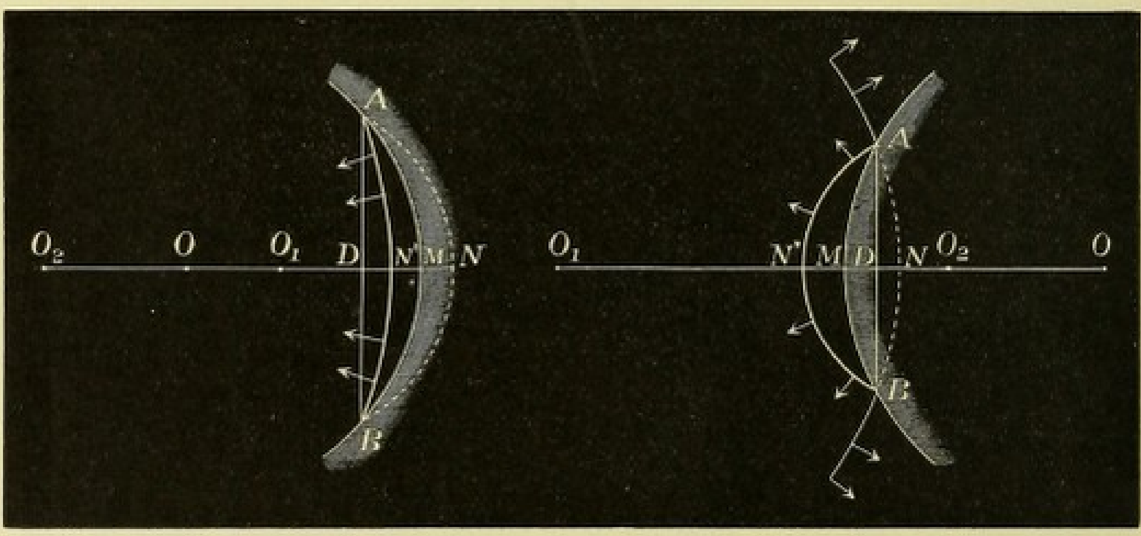



Fig. 39. Fig. 40.

spherical wave ANB (Fig. 39), diverging from a source $O_1$, to be reflected at the surface of a spherical mirror AMB, then the wave which would have occupied the position ANB, if unobstructed, will be converted into the wave AN′B by the reflecting action of the mirror, and the relation between the two waves is determined by the equality MN = MN′.

From this it follows at once that

$$DN + DN' = 2DM,$$

or the sum of the curvatures of the incident and reflected waves is equal to twice the curvature of the mirror.

In the case of a convex reflecting surface, if ANB (Fig. 40) represents the position which the incident wave would have gained if unobstructed by the reflecting surface, AN′B the reflected wave, then MN = MN′, and consequently

$$DN' - DN = 2DM.$$

But if we remark that in this case the curvature of the incident wave is opposite in sign to that of the reflected wave and of the mirror, we may write this equation like the foregoing in the form

$$DN + DN' = 2DM,$$

that is, with proper attention to sign we may say in general that the curvature of the reflected wave is equal to the curvature of the incident wave reversed added to twice the curvature of the mirror. This may also be expressed by saying that the curvature of the mirror is the arithmetical mean of the curvatures of the incident and reflected waves.

Denoting the curvatures of the incident and reflected waves by $\sigma_1$ and $\sigma_2$ respectively, and that of the mirror by $\sigma$, the fundamental equation for the reflection of a spherical wave at a spherical surface takes the form

$$\sigma_1 + \sigma_2 = 2\sigma.$$

Hence the sum of the curvatures of the incident and reflected waves is equal to the focal power of the mirror.

The interpretation of this equation is that if the incident wave diverges from a point $O_1$ at a distance $\rho_1$ from the mirror, the reflected wave will converge to (or diverge from) a point $O_2$ at a distance $\rho_2$ from the mirror such that

$$\frac{1}{\rho_1} + \frac{1}{\rho_2} = \frac{2}{\rho},$$

where $\rho$ is the radius of curvature of the mirror.

The points $O_1$ and $O_2$ are termed *conjugate foci* with regard to the mirror, and are such that a wave diverging from either will after reflection converge to (or diverge from) the other.

*Cor.*—When $O_1$ is at infinity the wave becomes plane—that is, $\rho_1$ is infinite, and the incident light forms a parallel beam, so that $\sigma_1 = 1/\rho_1 = 0$; and we have the equation of Art. 61, viz. $\sigma_2 = 2\sigma$ and $\rho_2 = \frac{1}{2}\rho = f$.

**63. Wave of any Form reflected at any Surface.** — In the preceding articles we have considered the reflection of plane and spherical waves, but the theory may be applied to a wave PQ (Fig. 41) of any form reflected at any surface AB. Consider the light incident

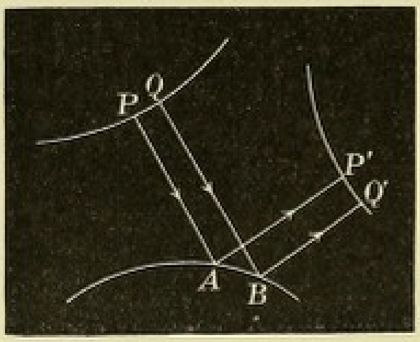



Fig. 41.

at any point A of the surface. This point is illuminated by the element of the wave at P where P is the pole of A, or that point of the wave from which it takes the light least time to reach A. If the curvature of the surface at A is not infinitely great the light will be reflected at A as from an element of the tangent plane at this point, so that the reflected light travels from A along a direction AP′, such that AP and AP′ make equal angles with the normal at A, and lie in the same plane with it according to the ordinary laws of reflection.

It may happen that the wave PQ has more than one pole with respect to A, that there may be several points P such that the time required by the light to reach A from them is either a maximum or a minimum. In this case the light will appear to come in rays from each of these points to A, and we will have a corresponding set of reflected rays. There might be a curve on PQ such that each point of it is a pole of A, the light then would appear to travel to A from the whole of this curve. Examples of such cases will appear in the sequel.

To find the form of the reflected wave take any system of points A, B, etc., on the reflecting surface and determine their poles P, Q, etc. With A, B, etc., as centres and radii $r$, $r'$, etc., such that $PA + r = BQ + r' =$ etc., describe spheres. These spheres touch the reflected wave P′Q′ at the points P′, Q′. The reflected wave may therefore be described either as the envelope of these spheres or as the locus of the points P′, etc., taken on the reflected rays such that $PA + AP' = QB + BQ' =$ constant. Hence if a set of rays be drawn perpendicular to any wave front in a homogeneous isotropic medium, they will after reflection (or any number of reflections) be perpendicular to the new wave front,[1] and the length of any ray from wave front to wave front will be constant and the same for all the rays. The same proposition holds likewise for refraction.

*Rays.*—From what has been proved we see that it is approximately legitimate to regard the light emitted from any point as made up of very narrow pencils or rays, and that after reflection each little beam

[1] Malus, *Journal de l'École Polytechnique*, cah. xiv. p. 1, 1808.

or ray is reflected on the other side of the normal to the surface at an angle equal to the angle of incidence. In dealing with problems in the reflection of light we may therefore consider the light propagated in rays if it facilitates the solution. Yet we must carefully bear in mind that *rays* have no physical existence, for it is waves that are propagated and not rays. The following examples are added for the sake of comparison :—

### *Examples*

1. If a plane mirror on which a pencil of light is incident be turned through any angle about an axis perpendicular to the plane of incidence, the reflected light is deviated through twice that angle.

[The direction of the incident light remains fixed in space, hence if the plane reflecting surface be turned through any angle $\theta$ the normal will be turned through the same angle, consequently the angle of incidence $i$ becomes $i \pm \theta$. The angle of reflection is also altered by the same quantity, therefore the angle between the incident and reflected rays is $2i \pm 2\theta$, but originally it was $2i$, therefore the reflected ray has been turned through an angle $2\theta$.

This theorem is of wide application in practice, for plane mirrors are extensively used to indicate, by the change in the direction of a reflected ray, the motions of magnetised or electrified needles, and for many other purposes in physical apparatus.]

2. Show from Ex. 1 that when a plane wave is reflected at a spherical surface the curvature of the reflected wave is equal to twice that of the surface.

3. Light emanating from a luminous origin O is reflected at a plane surface, prove from the doctrine of rays that the reflected light appears to come from a point O′ on the other side of the surface, such that the line OO′ is perpendicular to the surface, and O and O′ are equally distant from it (Fig. 42).

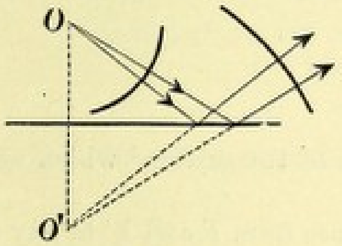



Fig. 42.

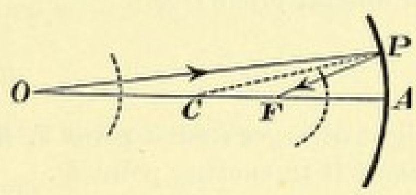



Fig. 43.

[The lines joining O and O′ to any point of the surface are equally inclined to it, hence every reflected ray passes through O′. An eye placed in the reflected light will receive a cone of light of which O′ is the vertex. The light will consequently appear to come from O′. This point is called the image, or reflection, of O in the surface. The reflected waves are spheres with O′ as centre.]

4. Light diverging from a point O is reflected at a concave spherical surface, find the conjugate focus by the doctrine of rays.

[Let C (Fig. 43) be the centre of the sphere and $\rho$ its radius, and let P be a point on the surface at a distance from A small compared with the distance OA or $\rho$. Then if OP be any incident ray and PF a reflected ray, CP is the normal to the surface, and therefore bisects the angle OPF. Hence

$$OP : PF :: OC : CF.$$

But approximately $OP = OA = \rho_1$, and $FP = FA = \rho_2$, therefore

$$\rho_1(\rho - \rho_2) = \rho_2(\rho_1 - \rho),$$

or

$$\frac{1}{\rho_1} + \frac{1}{\rho_2} = \frac{2}{\rho},$$

which is the algebraic statement that the row [AFCO] is harmonic—a property at once evident, for PA and PC are the bisectors of the angle OPF.

The reflected rays therefore pass through the point F determined by the above equation, and the reflected waves are spheres round F as centre. Of course this is only an approximation. When O is infinitely distant the light falls upon the mirror in a parallel beam, and the point F, to which it converges, is termed the *principal focus.* It lies halfway between C and A, or the principal focal distance $f$ is equal to $\frac{1}{2}\rho$. This follows from the above equation, for $1/\rho_1 = 0$ and $\rho_2 = f$. For a plane mirror $\rho = \infty$, so that $\rho_1 = -\rho_2$.]

5. The equation of Example 4 may be written in the form

$$(\rho_1 - f)(\rho_2 - f) = f^2,$$

where $f$ is the principal focal distance, or $2f = \rho$. $f$ is therefore a geometric mean between the distances of the conjugate foci O and F from the principal focus.

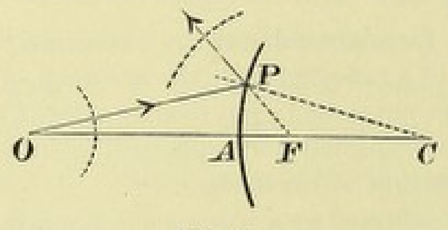


Fig. 44.

6. Light diverging from a point O falls upon a convex spherical mirror, find the position of the conjugate focus F.

[Here the point O (Fig. 44) and the centre of the mirror lie on opposite sides of the surface. The radius $\rho$ is therefore to be reckoned negative. Hence if $\rho_1$ is positive $\rho_2$ will be negative, and O and F will lie on opposite sides of the mirror. The focus F is in this case *virtual*, that is, the reflected light appears to diverge from it and the reflected wave is a sphere diverging from F as centre.]

7. If $p$ and $q$ denote the distances of two conjugate foci from the centre of a spherical mirror, prove that

$$\frac{1}{p} + \frac{1}{q} = \frac{2}{\rho}.$$

8. Light diverges from a point F, find the form of the surface[1] which will accurately reflect it to another point F′.

[The surface must evidently be such that the lines from F and F′ to any point of it are equally inclined to the tangent plane at that point. Now a fundamental property of an ellipse is that the lines joining the foci to any point of the curve are equally inclined to the tangent at that point. Hence light proceeding from one focus will be reflected to the other, and the surface generated by the revolution of an ellipse round its major axis will therefore satisfy the conditions of the problem. The surface is therefore any spheroid having F and F′ for foci. If the surface is a hyperboloid of revolution then one focus is the virtual image of the other, and if the surface is a paraboloid of revolution, the second focus being at infinity, the light proceeding from the first focus will after reflection travel in a parallel beam in the direction of the axis of the surface.

We arrive at the same conclusion by regarding conjugate foci as two points such that the time taken by light to travel from one to the other, by reflection at the

[1] Such a mirror is said to be *aplanatic.* A spherical surface is aplanatic for rays diverging from its centre only.

surface, is constant, or the same for all paths. For if $\rho$ and $\rho'$ be the distances of F and F′ from any point of the surface we have

$$\rho+\rho'=\text{constant},$$

but this is the fundamental property of an ellipsoid of revolution.

If mercury be placed in an elliptic dish and disturbed at one focus the reflected waves may be seen converging to the other focus.]

9. A luminous point is situated between two plane mirrors inclined at a given angle $\theta$, find the number and position of the images formed by successive reflections at the mirrors.

10. Show that when an eye is placed to view any image formed by successive reflections at two mirrors, the apparent distance of the image from the eye is equal to the distance actually travelled by the light in coming to the eye from the luminous point.

11. A luminous point is placed between two plane mirrors inclined at an angle of 27°. Prove that the number of images is thirteen or fourteen, according as the angular distance of the point from the nearer mirror is less than or greater than 9°.

12. If the light of the sun be admitted through a small hole an image of the sun is depicted on a screen placed to receive it, but if it be admitted through a large aperture we obtain an image of the aperture. Explain this.

[Each small portion of the aperture depicts an image of the sun, and the complete system of these images forms the image of the aperture.]

13. If the fraction of light reflected at the first surface of a parallel plate be $\alpha$ (there being no regular interference), that transmitted by the first surface, reflected by the second and again transmitted by the first, is $\alpha(1-\alpha)^2$. That reflected three times and transmitted twice is $\alpha^3(1-\alpha)^2$, etc. Hence the whole reflected light is

$$R=\alpha+(1-\alpha)^2(\alpha+\alpha^3+\alpha^5+\ldots)=\frac{2\alpha}{1+\alpha}.$$

14. The intensity of the light reflected from a pile of plates has been investigated by Provostaye and Desains (*Ann. de Chemie*, xxx. p. 159, 1850). If $\phi(m)$ be the reflection from $m$ plates the reflection from $m+1$ plates, as above, is

$$\phi(m+1)=\alpha+(1-\alpha)^2\phi(m)\{1+\alpha\phi(m)+\alpha^2(\phi(m))^2+, \text{ etc.}\}$$
$$=\frac{\alpha+(1-2\alpha)\phi(m)}{1-\alpha\phi(m)}.$$

But $\phi(1)=\alpha$, therefore we find $\phi(2)$, $\phi(3)$, etc., and generally

$$\phi(m)=\frac{m\alpha}{1+(m-1)\alpha}.$$

Stokes has extended this to the case in which the plates exercise an absorbing influence (*Proc. Roy. Soc.* xi. p. 545, 1862).

15. A system of rays being such that they all cut a given surface orthogonally, construct a mirror which will reflect the system to a given focus (Sir Wm. R. Hamilton, *Trans. Roy. Irish Academy*, vol. xv. p. 80, 1828).

[Take on each ray a point such that the sum (or difference) of its distances from the orthogonal surface and the given focus is constant. The locus of these points is the surface of the required mirror.]

16. If rays diverging from a point are reflected at any surface the reflected rays are cut orthogonally by a system of surfaces, and after reflection at any number of surfaces the whole length of each ray from the source to any orthogonal surface is the same for each ray (Hamilton, *ibid.*).

[The medium is supposed isotropic and the orthogonal surfaces are the wave surfaces.]

## CHAPTER V

### REFRACTION

**64. Refraction, Snell's Law.**—When light is incident on the surface of a transparent medium a portion is reflected; but another portion enters the medium, pursuing there in general an altered direction. This portion is said to be *refracted.* Generally we may say that when light is incident at the surface of separation of two media one portion is reflected back and propagated in the first medium, while another portion is refracted and transmitted through the second medium, if it be transparent, but absorbed immediately at the surface, or within a very small distance from it, if it be opaque.

The angle which the refracted ray makes with the normal to the surface is called the *angle of refraction.* If the first medium is rarer than the second, for example, air and water, the angle of refraction is less than the angle of incidence, the refracted ray is bent *towards* the normal (Fig. 45); on the other hand, it is bent or deviated *from* the normal if the first medium is optically denser—that is, more highly refracting than the second.

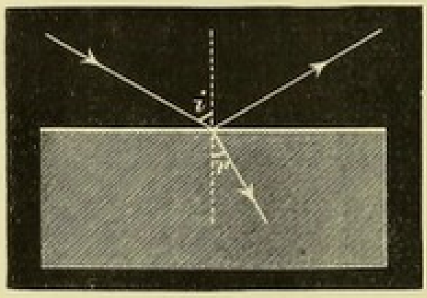


Fig. 45.

The relations connecting the angles of incidence and refraction are known as the *Laws of Refraction.* These were arrived at by Snell about 1621, but first stated by Descartes in the form—"The incident and refracted rays are in the same plane with the normal to the surface; they lie on opposite sides of it, and the sines of their inclinations to it bear a constant ratio to one another."

Denoting the angles of incidence and refraction by $i$ and $r$ respectively, the relation between them is stated in the formula

$$\frac{\sin i}{\sin r} = \mu.$$

The constant ratio $\mu$ is called the *index of refraction.* It is in general greater or less than unity according as the first medium is rarer or denser than the second.[1] When light passes from vacuum into any medium the ratio is termed its *absolute* index of refraction, but when it passes from one medium to another it is termed the *relative* index.

**65. Deduction of the Laws of Refraction.**—The theoretical deduction of the laws of refraction is in all respects similar to that of the laws of reflection. Let AB (Fig. 46), as before, be the trace of the incident plane wave, and AA′ that of the surface of separation, both planes being perpendicular to the plane of the paper. Let $v$ be the velocity of light in the first medium and $v'$ that in the second. Then if $t$ be the time required by the light to traverse the distance BA′, we have $BA' = vt = AD$, and if the wave had been unobstructed by the second medium, it would occupy the position A′ND (parallel to AMB) at the end of the time $t$. However, since the disturbance travels with a velocity $v'$ in the second medium, the point A becomes the centre of a spherical wave of radius $v't = AC$, or such that $AD : AC :: v : v'$, for $AD = vt$. Similarly, at the end of the time $t$ any point P of the surface will be the centre of a refracted wave of radius PM′, such that $PN : PM' :: v : v'$. Therefore

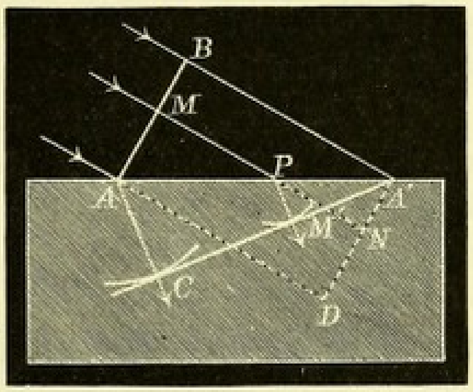


Fig. 46.

$$\frac{AC}{PM'} = \frac{AD}{PN} = \frac{AA'}{PA'},$$

and hence if a plane be drawn through A′ perpendicular to the plane of the paper to touch the sphere diverging from A, it will also touch that diverging from P, or from any other point of the surface AA′. The plane through A′C, perpendicular to the plane of the paper, is therefore the wave envelope or the locus of those points which are just about to be disturbed at the end of the time $t$. As in the case of reflection, we may show that the point M′ of the refracted wave is illuminated by the point P of the surface, that is, the disturbance at M is propagated along the path MPM′ from M to M′, and we shall show immediately that this path is that along which it takes light the least time to travel from M to M′, and that the principle of least time is obeyed in refraction as well as in reflection.

The angle AA′D is the angle of incidence and the angle AA′C is

[1] It is not universally true that the denser media are the more highly refracting, for example water, of unit density, has a refractive index 1·3336 for yellow light, whereas oil of turpentine, density 0·885, has an index 1·4744 for the same light.

the angle of refraction; denoting these by $i$ and $r$ respectively, we have

$$\frac{AD}{AC}=\frac{\sin i}{\sin r}=\frac{v}{v'}=\mu.$$

Hence we have the law of refraction, viz. "the sine of the angle of incidence bears a constant ratio to the sine of the angle of refraction," while we have also the additional information that this constant ratio, or the refractive index, is equal to the ratio which the velocity in the first medium bears to the velocity in the second.

The bending or deviation of a ray of light in passing from one medium to another is then due to the difference of the velocities of light in the two media. The greater the change of velocity the greater the bending. If light travels more slowly in the second medium than in the first, $v$ is greater than $v'$, $i$ is greater than $r$, $\mu>1$, and the refracted ray is deviated towards the normal; the reverse is the case when $v$ is less than $v'$.

Velocity test.

Now we know by direct observation that the deviation is towards the normal when light passes from a rare medium like air to a dense medium such as glass or water. The wave theory therefore indicates that the velocity of light in air is greater than its velocity in glass or water in the ratio of their refractive indices, and experiment proves this to be the case.

The emission theory, on the other hand, points to the opposite conclusion. According to it $\sin i/\sin r=v'/v$ (Art. 23), so that in those media where the bending is towards the normal the light travels with increased velocity. Here then the conclusions of the two theories are contradictory, and experiment, which alone can decide between them, supports the wave theory conclusively. The emission theory is consequently untenable in its original form, and requires serious modification in its fundamental tenets in order to meet this difficulty.

## *Examples*

1. If the velocities in two media be $v_1$ and $v_2$, while the velocity in vacuum is $v$, the *absolute* refractive indices of the media are

$$\mu_1=v/v_1, \text{ and } \mu_2=v/v_2,$$

while their *relative* refractive index, or the index of the second with respect to the first, is

$$\mu_{12}=v_1/v_2=\mu_2/\mu_1.$$

The law of refraction may consequently be written in the form

$$\mu_1 \sin i=\mu_2 \sin r.$$

2. For any number of media we have

$$\mu_{12} \cdot \mu_{23} \cdot \cdot \cdot \mu_{n-1,\ n} = \mu_n/\mu_1,$$

or the continued product of the relative refractive indices of $n$ substances is equal to the ratio of the absolute refractive index of the $n$th to that of the first.

**66. Construction for Reflected and Refracted Waves (Huygens).**—We have now the following construction for the wave fronts of the two portions into which a beam of parallel light is divided when it is incident on the surface of separation of two transparent isotropic media.

Let the plane of the paper be perpendicular to the plane of the incident wave, and also to the plane surface of separation of the media.

Let AB (Fig. 47) be the trace of the incident wave on the plane of

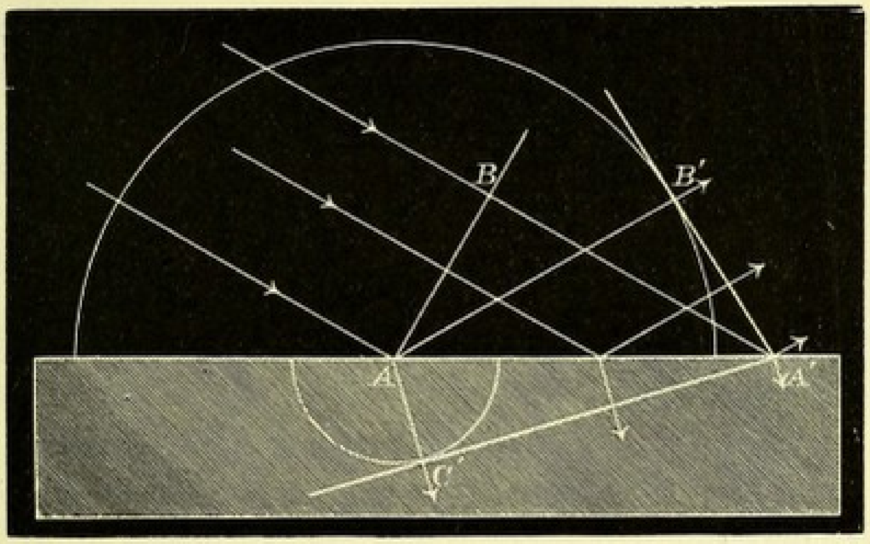



Fig. 47.

the paper. The plane of the wave is a plane through AB perpendicular to the plane of the paper, and the surface of separation is a plane through AA′ perpendicular to the paper. With A as centre describe two spheres of radii $vt$ and $v't$, where $t$ is the time of propagation from B to A′, and $v$ and $v'$ are the velocities of propagation in the first and second media respectively.

The radius AB′ of the upper sphere is equal to A′B, and the radius AC′ of the lower is A′B/$\mu$. The ratio of the radii is equal to the ratio of the velocities in the two media, or equal to the relative refractive index.

From A′ draw planes perpendicular to the paper to touch these spheres, and let A′B′ and A′C′ be the traces of these tangent planes.

These tangent planes are the limits to which the disturbance is propagated at the instant the wave from B reaches A′. They are the reflected and refracted wave fronts respectively. The perpendiculars

AB′ and AC′ on them from A are the reflected and refracted rays arising from the ray incident at A.

The media here considered are isotropic, and the waves diverging from any point are accordingly spherical. This is not the case in all media, yet the construction for the wave fronts remains the same: viz. with A as centre describe the waves, whatever shape they be, which have diverged from it at the instant the wave from B reaches A′. Through A′ draw tangent planes as before to these waves. These planes are the reflected and refracted wave fronts.

**67. Total Reflection.**—If the first medium be rarer (less refracting) than the second, the radius of the second sphere is less than that of the first, but the radius AB′ of the first is equal to BA′, which is less than AA′, hence the radius AC′ of the second sphere is always less than AA′, consequently A′ lies outside it, and it is always possible to draw a tangent plane to it from A′. There is then always a refracted wave. It is otherwise when the second medium is rarer (less refracting) than the first. In this case the velocity in the second is greater than in the first, and the radius of the second sphere is greater than the radius of the first. It is therefore greater than A′B, and may be greater than AA′ if the incidence exceeds a certain limiting value. If the radius of the second sphere is greater than AA′, the point A′ will lie inside it, and it will be impossible to draw a real tangent plane to it from A′. Consequently there is no refracted wave, and the light is all reflected back into the first medium, as shown in Fig. 48.

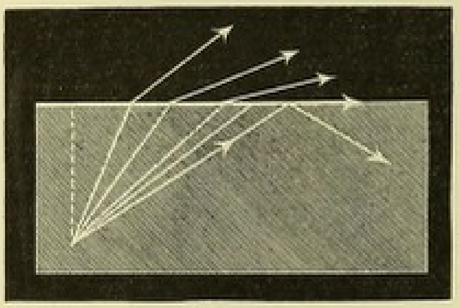


Fig. 48.

The light in this case is said to be totally reflected. This limit is reached when the radius AC′ of the lower sphere is equal to AA′. In this case

Critical angle.

$$\sin i = \frac{A'B}{A'A} = \frac{A'B}{AC'} = \frac{v}{v'} = \mu,$$

so that the limiting angle of incidence at which total reflection occurs is given by the equation

$$\sin i = \mu,$$

where $\mu$ is the refractive index of the rarer medium with respect to the denser. If, however, $\mu$ denotes the relative refractive index of the denser medium, the limiting angle is given by

$$\sin i = 1/\mu.$$

This angle is known as the *Critical Angle*, and total reflection occurs in the case of light passing into a less refracting medium if the angle of incidence is greater than the critical angle.

The existence of total reflection is frequently taken advantage of in the construction of optical instruments, and notice of it often comes within reach of ordinary observation, as when the surface of water is viewed in a glass held above the head, the silvery brilliancy of the surface being due to the total reflection of the light.

| | |
|---|---|
| For water the critical angle is about . . . . . . | 48° 27′ 40″ |
| For crown glass ,, ,, . . . . . . . . | 40° 30′ |
| For chromate of lead it is only . . . . . . . | 19° 28′ 20″ |

*Example*

If light is refracted at a plane surface, prove that the deviation—that is, the difference of the angles $i$ and $r$—increases as the angle of incidence increases.

[When $i$ is small we have $i=\mu r$, and therefore the deviation is $i-r=(\mu-1)r=i(\mu-1)/\mu$. The deviation consequently increases with the angle of incidence. When $i$ is not small it follows at once from the law of refraction that sin $i$ increases more rapidly than sin $r$, or that $i$ increases more rapidly than $r$—that is, that $i-r$ increases with $i$. This may be shown geometrically as follows:—Let OP (Fig. 49) be any line representing the velocity in the first medium, and on the same scale let OM represent the velocity in the second. With O as centre and OM as radius describe a circle. Then if the angle OPM be the angle of refraction the angle OMN will be the angle of incidence, for we have

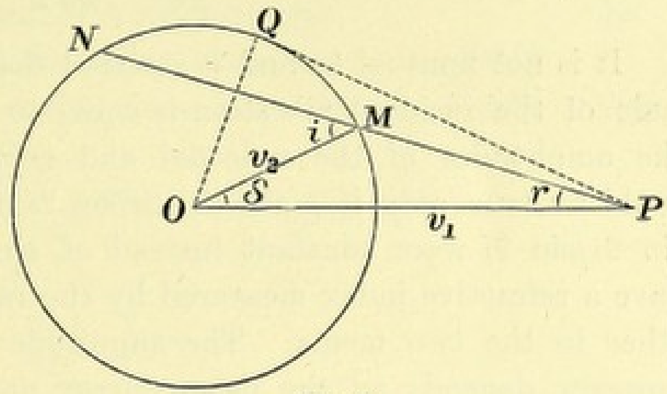


Fig. 49.

$$\frac{\text{OP}}{\text{OM}}=\frac{v_1}{v_2}=\mu=\frac{\sin i}{\sin r}=\frac{\sin \text{OMN}}{\sin \text{OPM}}.$$

Hence OMN $=i$, and if OP be parallel to the refracted ray and PM parallel to the normal then OM will be parallel to the incident ray. The angle POM is consequently equal to $i-r$ and therefore represents the deviation. Now it is clear that POM increases as $r$ increases—that is, as $i$ increases, until the line PM becomes a tangent to the circle. At this point, in the case of light refracted from the second medium into the first, the limit is reached beyond which total reflection occurs and refraction ceases.]

**68. Relation connecting the Intensities of the Incident, Reflected, and Refracted Waves—The Energy Equation.**—Since the incident wave is divided into a reflected wave and a refracted wave, its energy must be equal to the sum of the energies of the other two. The reflected and refracted waves derive their energies from the incident, and it is clear that a pencil of length $v$ of the latter gives rise to a reflected pencil of length $v$ and a refracted pencil of length $v'$, and the widths of the pencils are proportional to AB, A′B′, and A′C′ respectively (Fig. 47). Hence if $a$, $b$, $c$ be the amplitudes of the corresponding vibrations, the energy per unit volume will be proportional to $\rho a^2$, $\rho b^2$, $\rho' c^2$, where $\rho$ and $\rho'$ are symbols for the two media representing what we may call the density of the ether, or that property of it which corresponds to the density of ordinary matter, and by which it possesses energy when in motion. Hence the energy of an incident beam of length $v$ and width AB will be proportional to $v\rho a^2 \,.\, \mathrm{AB}$, and we have the equation

$$v\rho a^2 \,.\, \mathrm{AB} = v\rho b^2 \,.\, \mathrm{A'B'} + v'\rho' c^2 \,.\, \mathrm{A'C'}.$$

Hence

$$v\rho a^2 \cos i = v\rho b^2 \cos i + v'\rho' c^2 \cos r,$$

or, since $v/v' = \sin i/\sin r$, we have[1]

$$\frac{\rho(a^2-b^2)}{\rho' c^2} = \frac{\sin 2r}{\sin 2i}.$$

It is not unusual to find it asserted that the square of the amplitude of the incident vibration is equal to the sum of the squares of the amplitudes of the reflected and refracted vibrations; but this could be true only if $\rho/\rho' = \sin 2r/\sin 2i$, a law which might exist if $\sin 2r/\sin 2i$ were constant instead of $\sin i/\sin r$. We would then have a refractive index measured by the ratio of the densities of the ether in the two media. The amplitude of the refracted vibration, however, depends on the mean energy per unit volume, and this, we have seen, depends on the density $\rho'$ of the ether in the second medium or on the velocity of propagation.

*Cor.*—Two suppositions have been made with respect to the quantities $\rho$ and $\rho'$, one by Fresnel, that the velocity of propagation is inversely as the square root of the ether density, or that

$$\frac{\rho}{\rho'} = \frac{v'^2}{v^2} = \frac{\sin^2 r}{\sin^2 i},$$

by which the energy equation reduces to

$$\frac{a^2-b^2}{c^2} = \frac{\tan i}{\tan r}. \quad \text{(Fresnel's energy equation.)}$$

The other supposition, made by MacCullagh, is that $\rho = \rho'$, or the ether density is the same in all substances; we have then

$$\frac{a^2-b^2}{c^2} = \frac{\sin 2r}{\sin 2i}. \quad \text{(MacCullagh's energy equation.)}$$

[1] This equation may be written down at once by observing that the energy in the triangle ABA′ (Fig. 47) is equal to the sum of the energies in the triangles AB′A′ and AC′A′.

**69. The Principle of Least Time or the Law of Fermat.**—When light passes from any point M to another M′ (Fig. 31) by reflection at a surface, we have seen that the rays PM and PM′ are equally inclined to the surface, and consequently their sum is less than the sum of the lines joining M and M′ to any other point on the surface. The path MPM′ is that which will be traversed in the least time in passing from M to M′ by reflection at the surface.

A similar law governs the refraction of light, viz. if light pass from any point M (Fig. 50) to any point M′ in another medium, the path MAM′ traversed by the ray is such that the time occupied in travelling over it is a minimum. For if the time along the path MAM′ is a minimum, the time over this path must be equal to that occupied in traversing the consecutive (very near) path MA′M′. Hence if AB and A′C be drawn perpendicular to MA′ and M′A respectively, it follows that the times of travelling over the distances AC and A′B are equal, since MA = MB and M′A′ = M′C. Hence

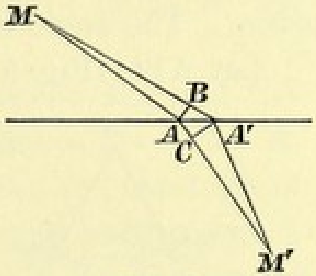


Fig. 50.

$$\frac{A'B}{v}=\frac{AC}{v'}, \quad \text{or} \quad \frac{AA'\sin i}{v}=\frac{AA'\sin r}{v'},$$

that is,

$$\frac{\sin i}{\sin r}=\frac{v}{v'}=\mu.$$

Hence if the time occupied in traversing the path be a minimum, the ordinary law of refraction is obeyed, and conversely.

Denoting the rectilinear paths in the two media by $l$ and $l'$, the law of Fermat asserts that

$$\frac{l}{v}+\frac{l'}{v'}=\text{a minimum},$$

or

$$l+\mu l'=\text{a minimum}.$$

If the same ray passes through several media we have $\Sigma(l/v)$, that is, $\Sigma\mu l$, a minimum, and if the refractive index of the medium changes from point to point, the path of a ray becomes a continuous curve, and the length $l$ in the above formula becomes a small element $ds$ of the curve. The principle of least time, according to which the wave front is always as far advanced as possible, may, for a medium of variable refractive index, be written in the form

$$\int\frac{ds}{v}=\text{minimum}, \quad \text{or} \quad \int\mu ds=\text{minimum}.$$

We are to observe, therefore, that in all cases of refraction through prisms, lenses, etc., when light travels from one point to another the ray pursues that path which requires least time. For example, when the various rays pass from one focus of a convex lens to its conjugate, those which travel through the centre of the lens transverse a greater distance in glass, and a less distance in air, than those which pass near the edge; but all rays require the same time to travel from one focus to the other, viz. the minimum time.

**70. Refraction of a Plane Wave through a Prism.**—A prism of any material is a wedge-shaped portion of it contained between two planes called its faces, which intersect in a line termed the edge of the prism. The angle between the faces is called the angle of the prism.

Let ABC (Fig. 51) be a section of a prism by the plane of the paper, supposed perpendicular to the edge of the prism. Consider a plane wave of light incident on the face AB in the direction PQ, the plane of incidence being the plane of the paper, and thus perpendicular to the edge of the prism. The light being refracted into the prism in the direction QR, making an angle $r$ with the normal, will suffer a deviation $i - r$ at the face AB. If the angle of incidence of QR on the second face be $r'$, and the angle of emergence along RS be $i'$, the light will suffer a further deviation $i' - r'$. The light will now emerge in the direction RS (if the angle of incidence $r'$ on the second face be less than the critical angle) and the total deviation $\delta$ from the original course PQ is

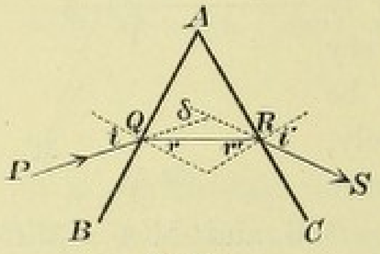


Fig. 51.

$$\delta = (i + i') - (r + r') = i + i' - A,$$

where A is the angle of the prism, which is equal to $r + r'$, since it is equal to the external angle between the normals at Q and R to the faces.

A plane wave BM (Fig. 52) incident on the face AB is refracted into the prism, and travels through it as a plane wave AD. It then emerges from the second face AC as a plane wave CN. Any point S of the wave CN is illuminated by a corresponding point P of the incident wave, and the law which governs the propagation of the disturbance is that the time of propagation along all paths joining pairs of corresponding points on BM and CN is the same. For example, the time along MAN = time along PQRS = time along BC. Hence we should have

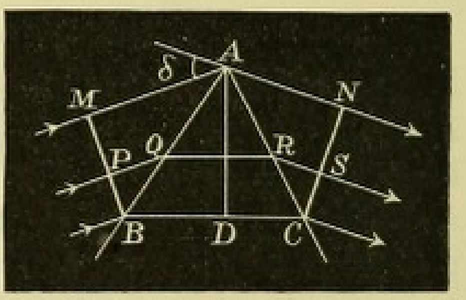


Fig. 52.

$$MA + AN = \mu BC = \mu(BD + DC),$$

or

$$AB \sin i + AC \sin i' = \mu(AB \sin r + AC \sin r'),$$

which is true if the law of refraction ($\sin i = \mu \sin r$) is obeyed.

Hence NC is such that the disturbances reach it from MB in the same time; they are therefore in the same phase, that is, NC is the wave front after refraction through the prism. The rays which pass near the edge of the prism traverse a shorter path in glass, but a longer path in air, than those which pass near the base, the long air path MAN occupying the same time as the shorter glass path BC, and the ratio of the lengths of these paths is the refractive index of the glass.

The deviation ($\delta$) is measured by the angle between MA and NA produced; but $MAB = 90 - i$, and $NAC = 90 - i'$; hence $\delta = i + i' - A$.

Now if $\mu$ be increased while the angle of incidence $i$ and the angle of the prism remain the same, the angle of refraction $r$ will be diminished, for $\sin i = \mu \sin r$. Hence the angle of incidence $r'$ on the second face will be increased, for $r + r' = A$, and consequently the angle of emergence $i'$ will also be increased to obey the relation $\sin i' = \mu \sin r'$. Hence if, while $i$ and A remain the same, the refractive index be increased, the angle of emergence $i'$ will be increased, and the total deviation $i + i' - A$ will be increased by the same amount.

The amount of deviation, therefore, depends on the refractive index, that is, upon the velocity of propagation in the prism. We should consequently expect that if ordinary solar light contains many constituent waves which travel with different velocities in the prism, they should be deviated by different amounts on emerging, and a beam of parallel light incident on the first face should be a dispersed beam on emergence from the second face. That this is the case was first demonstrated by Newton.[1] Dispersion.

**71. Colour and Velocity.**—The experiments of Newton prove in a clear and masterly manner that lights of different colours are refracted by different amounts in passing through a glass prism or on entering a new medium, the red light being deviated least, the violet most, and the intermediate colours (orange, yellow, green, and blue) in continuous transition by intermediate amounts. But we have seen that the bending of the ray on entering a new medium is a consequence of the difference of velocity in the two media, and the greater this difference the greater the deviation.

We therefore conclude that the velocity of the violet light is least and that of the red greatest in the prism, while the intermediate

[1] See extracts at end of chapter.

colours travel with intermediate velocities. We have reason to believe that all the colours travel with the same velocity in free space, and with practically the same velocity in air, for if the red travelled faster than the violet it would follow that a star reappearing after eclipse should at first appear red, as the red light would reach us first, and then gradually change tint till it finally became white when all the colours have had time to arrive. Similarly, when the star is just disappearing at eclipse it should be violet coloured, as the violet would be the last to reach us.

Now as no observation of this nature has ever been made it follows that the violet waves must differ in some respect from the red, and this difference must exist either in the wave length or the periodic time of vibration, or both, for if they all had the same periodic time and wave length there would be nothing left by which we could distinguish the red waves from the violet. The waves then in air must have different periodic times and different lengths, for since they travel with the same velocity, the equation

$$v\mathrm{T}=\lambda$$

shows that the periodic times are proportional to the wave lengths.

In the case of refraction there is one element which is likely to remain unaltered, viz. the periodic time of vibration. For the vibration in the second medium is excited and forced by the vibration in the first medium, and these will in general be executed in the same time. The time T then is the same in the incident and refracted wave, so that if the velocity changes in the second medium the wave length changes proportionately by the equation

$$v'\mathrm{T}=\lambda'.$$

Hence it follows that when refraction occurs

$$\frac{v}{v'}=\frac{\lambda}{\lambda'}=\mu,$$

and the quantities

$$\frac{\lambda}{v}=\frac{\lambda'}{v'}=\mathrm{T}$$

remain unaltered.

So far as we have yet gone the theory has not indicated whether the red waves are longer or shorter than the violet; all we have arrived at is that the shorter the wave length the shorter the periodic time is in a vacuum. Further on experiments will be given which show that the red waves execute about 395 billion ($395 \times 10^{12}$) vibrations per second, while the violet vibrate about 763 billion ($763 \times 10^{12}$) times per second, or nearly twice as fast as the red.

The dispersion of the colours (as it is called) by a prism shows that although all the waves travel with the same velocity in free space, yet in dense media, like glass and water, the velocity of propagation depends upon the wave length or periodic time.

*Exercise.*—Light incident at an angle $i$ is refracted through a parallel plate of glass. Waves of lengths $\lambda_1$ and $\lambda_2$ are refracted at angles $r_1$ and $r_2$ respectively, find the relative retardation by transmission through the plate.

Let PQ (Fig. 53) be the front of a plane wave incident on the face of the plate at an angle $i$, and QO the direction of the incident light. Let PM and PN be the directions of the refracted rays. The planes $OMR_1$ and $ONR_2$, perpendicular to PM and PN, are the refracted wave fronts, which on emergence, if the plate be parallel, will proceed parallel to PQ. Thus the emergent wave fronts are $R_1B$ and $R_2C$, and if the plate had not been present the original wave front would have been propagated to OLA. Hence the effect of the plate is to retard one wave by an amount AB and the other by $AR_2$. The relative retardation of one on the other is therefore $R_2B$.

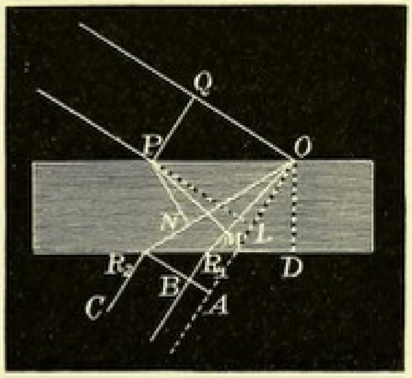



Fig. 53.

But if $e$ be the thickness of the plate, we have $R_1D = e \cot r_1$, $R_2D = e \cot r_2$, hence

$$
\begin{aligned}
R_2B = R_1R_2 \sin i &= \sin i(R_2D - R_1D),\\
&= e \sin i(\cot r_2 - \cot r_1),\\
&= e(\mu_2 \cos r_2 - \mu_1 \cos r_1).
\end{aligned}
$$

This example will be of use hereafter in the consideration of the colours of mixed plates and of doubly refracting crystals.

*Cor.*—When $i = r = 0$, the light is incident normally, and we have the retardation $= e(\mu_2 - \mu_1)$, as is otherwise directly obvious.

**72. Refraction of a Spherical Wave at a Plane Surface.**—Let spherical waves diverging from a centre O (Fig. 54) fall upon the plane surface AB of a transparent substance. If the wave had not encountered the second medium it would at some instant occupy the position ANB, a sphere of radius OA and centre O. The velocity in the second medium is, however, different from that in the first, so that M has become the centre of a wave in the second medium of radius MN′, where $MN : MN' = v_1 : v_2$, or

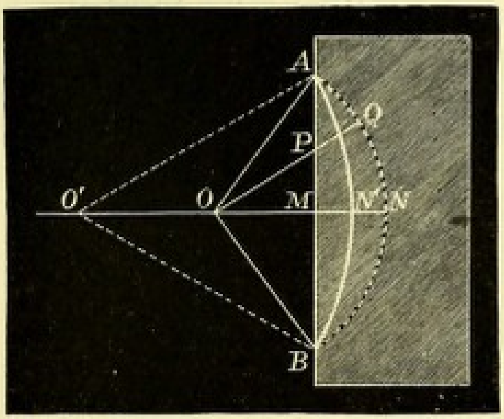



Fig. 54.

$$MN = \mu \,.\, MN'.$$

Similarly any other point P will be the centre of a spherical wave of radius equal to $PQ/\mu$, and the refracted wave will be the envelope of all these spheres. In the figure we have supposed the second substance more refracting than the first, so that their relative index $\mu$ is greater than unity. The consequence is that the incident wave ANB is flattened down into another AN′B of less curvature.[1]

Now MN and MN′ are proportional to the curvatures of the two waves. Hence the relation between the curvatures is determined by the foregoing equation, and may be expressed by saying that the curvature of the incident wave is $\mu$ times that of the refracted wave. Denoting these curvatures by $\sigma_1$ and $\sigma_2$ respectively, this relation may be represented in any one of the following forms :—

$$\sigma_1 = \mu\sigma_2, \quad \text{or } \sigma_1/\sigma_2 = v_1/v_2, \quad \text{or } \mu_1\sigma_1 = \mu_2\sigma_2,$$

where $\mu_1$ and $\mu_2$ are the absolute indices of the first medium and second medium respectively.

If $\rho_1$ and $\rho_2$ represent the radii of curvature, we have $\sigma_1 = 1/\rho_1$, $\sigma_2 = 1/\rho_2$, and consequently the relation between the distances of the points O and O′ from the surface is

$$\rho_2 = \mu\rho_1, \quad \text{or } \rho_1 v_1 = \rho_2 v_2, \quad \text{or } \rho_1/\rho_2 = \mu_1/\mu_2.$$

**73. Refraction of a Plane Wave at a Spherical Surface.**— The case of a plane wave incident on a spherical refracting surface is one of extreme simplicity when considered from the point of view of the wave theory. Thus let AMB (Fig. 55) represent the surface of a refracting sphere and XABY the trace of a plane parallel to the front of the incident wave. If the refracting sphere had not been present XY would represent the incident wave in one of its positions, but by the action of the sphere the portion ANB of the plane wave is retarded and takes the form of the curve AN′B, which is approximately an arc of a circle when AB is small. Now MN and NN′ are proportional to the curvatures of the surface and refracted wave respectively, and the relation between these is determined by the equation

$$\text{MN} = \mu\text{MN}' = \mu(\text{MN} - \text{NN}').$$

That is

$$\mu\text{NN}' = (\mu - 1)\text{MN}.$$

Hence if $\sigma$ be taken to represent the curvature of the surface and $\sigma'$ that of the refracted wave, we have

$$\sigma' = \frac{\mu - 1}{\mu}\sigma.$$

[1] If AB is small the arc AN′B will be approximately a circle having its centre at O′. The true form of the refracted wave is not a sphere, however, but a parallel to a hyperboloid, so that the curve AN′B is a parallel to a hyperbola (see Ex. 1, p. 111).

The case of concave refracting sphere is shown in Fig. 56, where N′N′ is the front of the refracted wave and NN is the corresponding position of the incident wave. Here again we have $MN = \mu MN'$, from which we derive the same formula as that obtained for a convex surface. In both cases therefore the action of the refracting sphere is to imprint on the refracted wave a curvature equal to $(\mu - 1)/\mu$ times that of the refracting sphere. This impressed curvature $\sigma'$ is always less than that of the refracting surface $\sigma$, and when $\mu$ is greater than unity—that is, when the sphere is more highly refracting than the medium in which it is immersed—the curvature of the refracted wave is of the same sign as that of the surface, but of opposite sign when $\mu$ is less than unity. Hence when $\mu$ is greater than unity the focus F, Impressed curvature.

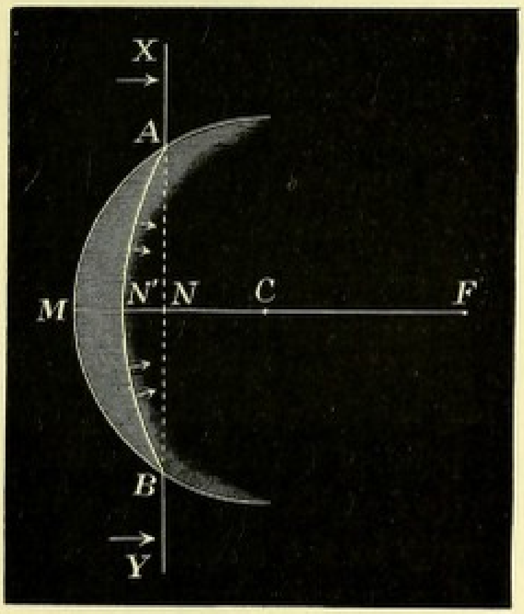



Fig. 55.

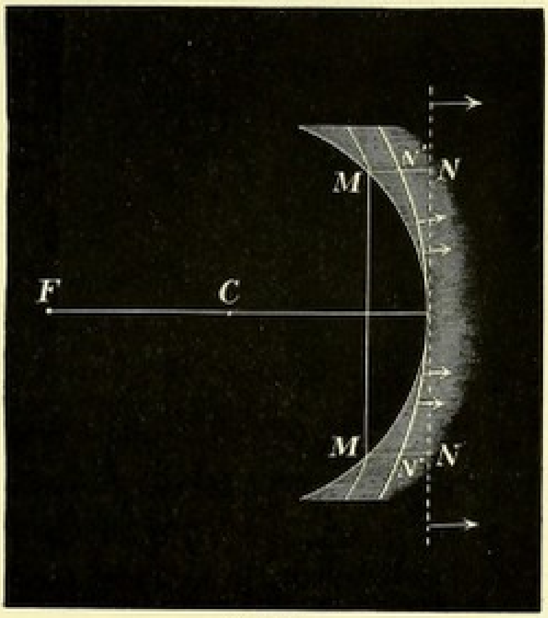



Fig. 56.

to which the refracted wave converges (or from which it appears to diverge), lies on the same side of the surface as its centre C; but when $\mu$ is less than unity the focus and centre lie on opposite sides of the surface. The relation between the focal distance $f = FM$ and the radius $\rho$ of the refracting surface is, since $\sigma' = 1/f$ and $\sigma = 1/\rho$,

$$f = \frac{\mu}{\mu - 1}\rho.$$

If the absolute indices $\mu_1$ and $\mu_2$ be used instead of the relative index $\mu = \mu_2/\mu_1$ the above formulæ become

$$\sigma' = \frac{\mu_2 - \mu_1}{\mu_2}\sigma, \quad \text{and } f = \frac{\mu_2}{\mu_2 - \mu_1}\rho.$$

The curvature $\sigma'$ impressed on a plane wave by refraction at a spherical surface is called the *focal power* of the surface. It is measured by the Focal power.

reciprocal of $f$, and means the curvature impressing power of the surface.

It has now been proved that the refraction of a spherical wave at a plane surface changes its curvature from $\sigma_1$ to $\sigma_2 = \sigma_1/\mu$, while the refraction of a plane wave at a spherical surface impresses a curvature $\sigma(\mu - 1)/\mu$ on the refracted wave, and we shall see in the following article that when a spherical wave is refracted at a spherical surface the curvature of the refracted wave is the sum of these two quantities.

**74. Refraction of a Spherical Wave at a Spherical Surface.**— Let the surface AB (Fig. 57) of the refracting medium be a sphere of centre C, and let a spherical wave AN′B diverging from O meet it at A and B. If the second substance be more refracting than the first, the spherical wave, which at any instant would have occupied the position ANB, will be flattened into the surface AN′B. This surface is the envelope of the sphere described with any point P of the surface as centre and radius PQ/$\mu$. It is the wave surface, and such that the disturbances from O reach every point of it in the same time, viz. the interval required to travel directly from O to A or B. The refracted wave AN′B is propagated normally in the second medium, and if ONN′ be a common normal to both waves, in any position we have

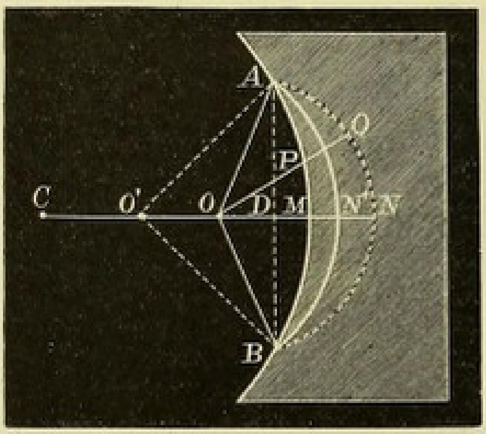


Fig. 57.

$$MN = \mu MN'.$$

The effect of the refraction is therefore to diminish or increase the curvature of the wave surface according as the velocity is less or greater in the second medium.

In the same manner we may construct the refracted wave when the incident wave and the refracting surface have any form.

If the angle AOB be very small the refracted wave will be approximately a sphere diverging from a centre O′, and the relation connecting the curvatures of the two waves with that of the surface may be easily determined, for the equation $MN = \mu MN'$ may be written in the form

$$DN - DM = \mu(DN' - DM).$$

That is

$$DN = \mu DN' - (\mu - 1)DM.$$

Hence if the curvatures of the incident wave, refracted wave, and surface be denoted by $\sigma_1$, $\sigma_2$, and $\sigma$ respectively, we have the relation

$$\sigma_1 = \mu\sigma_2 - (\mu - 1)\sigma, \quad \text{or } \frac{\sigma - \sigma_1}{\sigma - \sigma_2} = \mu.$$

Therefore $\sigma_2$ is determined by the equation

$$\mu\sigma_2 - \sigma_1 = (\mu - 1)\sigma, \quad \text{or } \sigma_2 = \frac{\sigma_1}{\mu} + \frac{\mu - 1}{\mu}\sigma.$$

The relation between the conjugate focal distances is consequently

$$\frac{\mu}{\rho_2} - \frac{1}{\rho_1} = \frac{\mu - 1}{\rho}.$$

If the relative index $\mu$ be replaced by $\mu_2/\mu_1$, the ratio of the absolute indices of the two media, the relation between the curvatures takes the symmetrical form

$$\mu_1\sigma_1 - \mu_2\sigma_2 = (\mu_1 - \mu_2)\sigma,$$

and the relation between the conjugate focal distances becomes

$$\frac{\mu_1}{\rho_1} - \frac{\mu_2}{\rho_2} = \frac{\mu_1 - \mu_2}{\rho}.$$

The same formula holds good for a convex refracting surface if the sign of the curvature of the incident wave be reckoned opposite to that of the surface.

*Cor.*—If $p$ and $q$ denote the distances of the conjugate foci O and O′ (Fig. 57) from the centre C of the refracting surface, then $\rho_1 = \rho - p$, $\rho_2 = \rho - q$, and the foregoing formula becomes

$$\frac{\mu_1}{q} - \frac{\mu_2}{p} = \frac{\mu_1 - \mu_2}{\rho},$$

or in terms of the relative index $\mu$ we have

$$\frac{\mu}{p} - \frac{1}{q} = \frac{\mu - 1}{\rho}.$$

### *Example*

A pencil of light diverging from a point falls directly on a refracting sphere of radius $\rho$ and passes through it, find the focus of the transmitted rays.

**75. Refraction through a Lens.**—A lens is a portion of a transparent substance, such as glass, quartz, or rock-salt, bounded by two surfaces of such a shape that light diverging from a point O (Fig. 58) and falling on one face, will after transmission converge to, or diverge from, another point O′. That is, the surfaces of the lens are so shaped that a wave which is spherical before transmission remains spherical after transmission. The surfaces of the lens itself are not accurately spherical, but only approximately so. The true curve of the surface which will accurately refract to a point light diverging from another point is the Cartesian oval (see Ex. 7, p. 113). A surface which refracts to a point the light diverging from another point is called an *aplanatic surface*, and the points are called *conjugate foci* with respect to it. The fundamental relation connecting conjugate foci O and O′ is that the time of propagation is the same along all the paths by which the light reaches O′ from O. Thus in Fig. 58 if light diverging from O is refracted by a double convex lens AB to the point O′, any spherical wave diverging from O emerges from the lens as a spherical wave con-

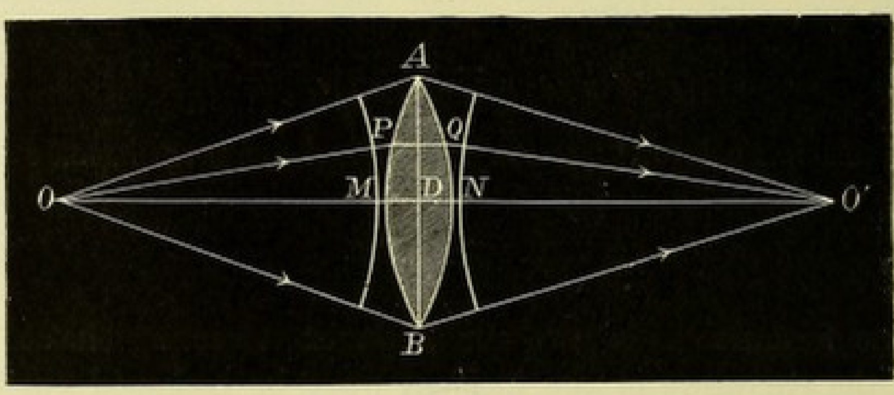



Fig. 58.—Aplanatic Lens.

verging to O′. Every point of this wave is a pole with respect to O′, and all the disturbances which simultaneously reach O′ are in the same phase. The illumination at O′ is consequently very intense. At points outside the cone O′AB there is destructive interference and darkness.

The relation connecting the conjugate focal distances, or the curvatures of a wave before and after transmission through a lens, follows very simply from the principles of the wave theory. In the first place, let us consider the case of a concavo-convex lens (Fig. 59) in which the thickness MN and the aperture AB are small compared with the focal distances. This is taken as a typical case because the curvatures of its surfaces are of the same sign, whereas in the double convex lens (Fig. 58) the curvatures of the faces are of opposite signs.

The relation which we are about to deduce might be obtained immediately by applying the formula of Art. 74 to the two faces of the lens in succession, but for the sake of illustration we shall approach the problem directly from fundamental principles. The principle on which the present investigation is based is that the time required by light to traverse the path OAO′ is the same as that required for the path OMNO′. The air equivalent of the latter path is OM + $\mu$MN + NO′. Hence the fundamental equation is

$$OA + O'A = OM + O'N + \mu MN.$$

With centre O and radius OA describe a circle cutting OO′ at P, and with centre O′ and radius O′A describe a circle cutting OO′ at Q, then, writing the foregoing equation in the form

$$(OA - OM) + (O'A - O'N) = \mu MN,$$

we have

$$- PM + QN = \mu MN,$$

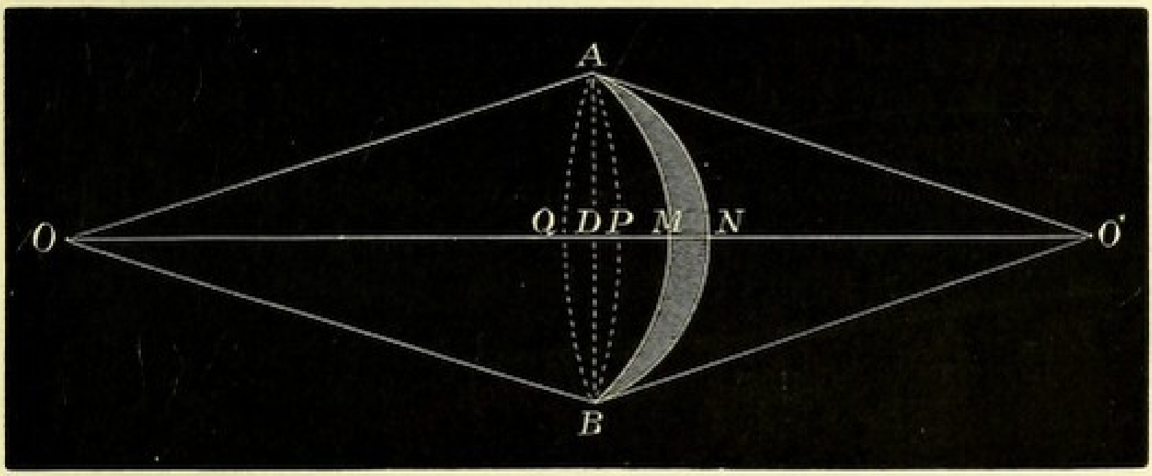



Fig. 59.

which, expressed in terms of the sagittæ, gives at once

$$DP - DM + DQ + DN = \mu(DN - DM).$$

Consequently the relation connecting the curvatures of the waves entering and emerging from the lens with the curvatures of its faces is

$$DP + DQ = (\mu - 1)(DN - DM),$$

or denoting the curvatures of the waves as before by $\sigma_1$ and $\sigma_2$, and the curvatures of the surfaces by $\sigma$ and $\sigma'$, we have

$$\sigma_1 + \sigma_2 = (\mu - 1)(\sigma - \sigma')$$

where $\sigma$ refers to the surface of greater curvature, which in Fig. 59 is the second: that on which the light is incident being regarded as the first surface of the lens.

When expressed in terms of the corresponding radii of curvature the foregoing equation furnishes the relation between the conjugate focal distances, viz.—

$$\frac{1}{\rho_1} + \frac{1}{\rho_2} = (\mu - 1)\left(\frac{1}{\rho} - \frac{1}{\rho'}\right),$$

where $\rho_1$ and $\rho_2$ are the distances of O and O′ from the lens, and $\rho$ and $\rho'$ are the radii of curvature of its faces ANB and AMB respectively.

If the point O from which the light emanates be infinitely distant the incident wave will be plane—that is, its curvature $\sigma_1$ will be zero, and the curvature of the transmitted wave will be

$$\sigma_2=(\mu-1)(\sigma-\sigma').$$

Focal power of a lens.

The function of the lens is consequently to change the curvature of a wave by an amount $(\mu-1)(\sigma-\sigma')$, so that this quantity represents the curvature producing power and is termed the *focal power* of the lens. In general, therefore, we may say that the curvature of the transmitted wave is equal to the algebraic sum of the curvature of the incident wave and the focal power of the lens—that is, the final curvature of the wave is equal to the algebraic sum of its initial and impressed curvatures.

The point to which a plane wave or a parallel beam of light is concentrated is termed the *principal focus* of the lens, and its distance from the lens is termed the focal length. Denoting the focal length by $f$ we have

$$\frac{1}{f}=(\mu-1)\left(\frac{1}{\rho}-\frac{1}{\rho'}\right),$$

which expresses the focal power in terms of the refractive index and the curvatures of the faces of the lens.

In this calculation the thickness of the lens has been supposed to be greatest at the centre. In some lenses, however, the thickness increases from the centre towards the edge, so that their thinnest part is at the centre. Lenses are accordingly divided into two classes according as their greatest or least thickness is at the centre. The former are termed *convex* and the latter *concave* lenses.

An examination of the motion of O′ as O moves along the axis of the lens will be a useful exercise, and the deduction of the formulæ for the various forms of lenses presents no further difficulty.

## *Examples*

1. Prove that the distance between two conjugate foci with respect to a lens cannot be less than four times its focal length.

[Since the sum of the reciprocals of two conjugate distances is equal to $1/f$, and therefore constant for a given lens, it follows that the product of these reciprocals is greatest when they are equal, and consequently the sum of the distances themselves must be least when they are equal. Thus

$$\frac{1}{f^2}=\left(\frac{1}{\rho_1}+\frac{1}{\rho_2}\right)^2=\left(\frac{1}{\rho_1}-\frac{1}{\rho_2}\right)^2+\frac{4}{\rho_1\rho_2},$$

which shows that the product of the reciprocals is greatest (*i.e.* $\rho_1\rho_2$ is least) when $\rho_1=\rho_2$. But $\frac{1}{f}=\frac{1}{\rho_1}+\frac{1}{\rho_2}=\frac{\rho_1+\rho_2}{\rho_1\rho_2}$.

Hence $\rho_1+\rho_2$ must bear a constant ratio to $\rho_1\rho_2$, and consequently $\rho_1+\rho_2$ is least when $\rho_1\rho_2$ is least, which is when $\rho_1=\rho_2=2f$. In other words, the least distance between two conjugate points is $4f$.]

2. A lens is employed to cast an image of an object on a screen, show that in general there are two possible positions of the lens for given positions of the object and screen, and prove that the length of the object is a geometrical mean between the lengths of the two images.

[In order that any image may be possible the distance between the object and screen must be greater than $4f$. This being the case, let the lens produce a distinct image on the screen when placed at a distance $\rho_1$ from the object and $\rho_2$ from the screen. Then if $l$ be the length of the object and $l_1$ the length of the image, it follows that

$$\frac{l}{l_1}=\frac{\rho_1}{\rho_2} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (1)$$

Now since conjugate foci are interchangeable it follows that the lens will also produce a distinct image on the screen when the distance of the lens from the object is $\rho_2$ and from the screen $\rho_1$, so that if $l_2$ be the length of the image in this case we have

$$\frac{l}{l_2}=\frac{\rho_2}{\rho_1} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (2)$$

Multiplying (1) and (2) together we have

$$l^2=l_1l_2, \text{ or } l=\sqrt{l_1l_2}.$$

Hence if $l_1$ and $l_2$ be measured, the length $l$ can be determined when the object is not accessible for the purpose of direct measurement.

This method of obtaining the length of a luminous object, or the distance between two luminous points is of practical importance in measurements of the wave length of light (p. 144).]

3. If a lens of absolute index $\mu_3$ be situated between two media of absolute indices $\mu_1$ and $\mu_2$, prove that the curvatures $\sigma_1$ and $\sigma_2$ of the waves entering and emerging from the lens are connected with the curvatures $\sigma$ and $\sigma'$ of the two faces of the lens by the equation

$$\mu_1\sigma_1+\mu_2\sigma_2=(\mu_1-\mu_3)\sigma-(\mu_2-\mu_3)\sigma'.$$

[In this case we have, by the principle of least time (Fig. 59),

$$\mu_1\text{OA}+\mu_2\text{O}'\text{A}=\mu_1\text{OM}+\mu_3\text{MN}+\mu_2\text{O}'\text{N},$$

which may be reduced at once, as in Art. 75.]

4. Find the relation connecting the curvatures of the incident and emergent waves in the case of a lens of which the thickness $c$ is not small.

[As before let $\sigma_1$ be the curvature of the incident wave when it reaches the first surface $\sigma$ of the lens, then the curvature $\sigma'_2$ of the wave when it just enters this surface is given by the equation (Art. 74)

$$\mu_1\sigma_1-\mu_2\sigma'_2=(\mu_1-\mu_2)\sigma \quad . \quad . \quad . \quad . \quad . \quad . \quad (1)$$

Now the wave when it enters the lens travels a distance $c$ before it meets the second surface, and consequently its curvature $\sigma''_2$ when it meets the second surface is by the Ex. of Art. 60

$$\sigma''_2=\frac{\sigma'_2}{1+c\sigma'_2} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (2)$$

But by Art. 74 the curvature $\sigma_2$ of the wave emerging from the lens is connected with $\sigma'$, the curvature of the second surface and with $\sigma''_2$ by the equation

$$\mu_2\sigma''_2 - \mu_1\sigma_2 = (\mu_2 - \mu_1)\sigma',$$

and this by (2) becomes

$$\frac{\mu_2\sigma'_2}{1 + c\sigma'_2} - \mu_1\sigma_2 = (\mu_2 - \mu_1)\sigma'.$$

Substituting for $\mu_2\sigma'_2$ from (1) we have at once

$$\mu_1\sigma_2 = \frac{\mu_1\sigma_1 + (\mu_2 - \mu_1)\sigma}{1 + \{\mu_1\sigma_1 + (\mu_2 - \mu_1)\sigma\}c/\mu_2} - (\mu_2 - \mu_1)\sigma'.$$

This equation connects the distances, $\rho_1 = 1/\sigma_1$ and $\rho_2 = 1/\sigma_2$, of any pair of conjugate foci from the surfaces of the lens with the radii of curvature of its faces and the thickness $c$. If the incident wave be plane we have $\sigma_1 = 0$, and the focal power of the lens is given by the equation

$$\frac{1}{f} = \sigma_2 = \left(\frac{\mu_2 - \mu_1}{\mu_1}\right)\left\{\frac{\sigma}{1 + (\mu_2 - \mu_1)\sigma c/\mu_2} - \sigma'\right\}.$$

This expression shows that the focal power is not the same when the light falls on the surface $\sigma'$ as when it falls on $\sigma$, so that the focal power changes when the lens is reversed with regard to the incident light. When the thickness is neglected this reduces to the expression of Art. 75, and the focal length remains unaltered when the lens is reversed.]

5. Determine the focal power of a combination of two thin lenses situated at a distance $c$ apart.

[Let $f_1$ and $f_2$ be the focal lengths of the two lenses, then if a plane wave falls upon the lens $f_1$, its curvature when emerging from it is $1/f_1$, and its curvature when it reaches the second lens is $\frac{1}{f_1 + c}$. Hence the curvature of the wave emerging from the second lens is

$$\frac{1}{f_1 + c} + \frac{1}{f_2},$$

which expresses the equivalent focal power of the combination. Denoting this by $\phi$ and denoting the focal powers of the lenses by $\phi_1$ and $\phi_2$ respectively we have

$$\phi = \frac{\phi_1}{1 + c\phi_1} + \phi_2.$$

The focal power when the combination is reversed is obtained by interchanging $\phi_1$ and $\phi_2$ or $f_1$ and $f_2$ in the foregoing expressions.

If a wave of curvature $\sigma$ falls upon the combination its curvature after transmission will be $\sigma + \phi$.]

## CAUSTICS

**76. Evolute of Wave.**—The reflected and refracted waves which we have so far considered have been either plane or spherical in form, so that the reflected and refracted rays converge to, or diverge from, a point. In the case of plane waves this point is infinitely distant.

In general, however, when light diverging from a point is reflected or refracted at a given surface, the front of the reflected or refracted wave is neither spherical nor plane, and the rays will not pass through a single point, but will envelop a surface called the *caustic*. This want of convergence of the rays is termed *astigmatism*.

Now the rays are normals to the wave, and the normals to a curve (or surface) envelop another curve (or surface) called the evolute.[1] Consequently the reflected or refracted rays envelop the evolute of the reflected or refracted wave. The evolute of the reflected wave is therefore the *caustic by reflection*, and the evolute of the refracted wave is the *caustic by refraction*.

It will be sufficient for our present purpose to confine our attention to surfaces of revolution, the incident light diverging from a point on the axis of revolution. In this case the caustics will also be surfaces of revolution about the same axis, and the section of the caustic surface by any plane through the axis will be the caustic curve of the generating curve of the surface at which the reflection or refraction occurs.

Thus if a surface be generated by the revolution of the curve APQ (Fig. 60) round the axis OA, and if O be the origin of light, then rays OP, OQ falling upon the surface will be refracted in directions PM and QN. The curve BMN cutting these refracted rays at right angles, or rather the surface formed by the revolution of BMN around OB, will be the refracted wave. The evolute of the curve BMN is the envelope of the rays refracted at APQ, that is, the caustic surface is formed by the revolution of the caustic curve around OA.

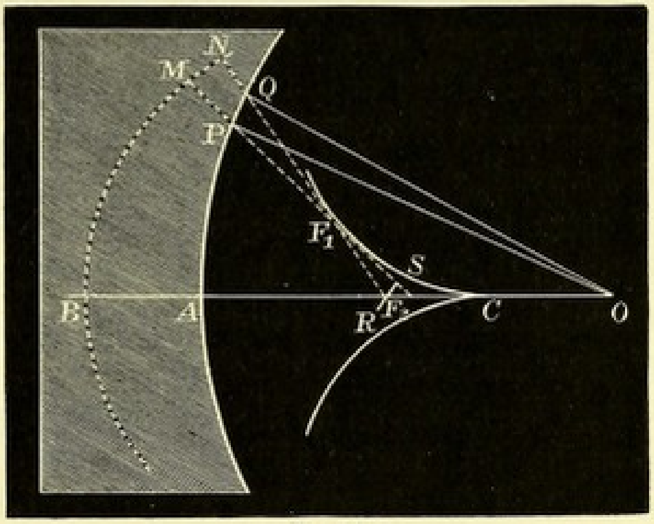


Fig. 60.

**77. Primary and Secondary Foci.**—Let the refracted rays PM and QN, produced backwards, intersect at $F_1$; then $F_1$ is a point on the caustic curve if the rays PM and QN be considered infinitely near.

[1] In the general case when the wave is not a surface of revolution the caustic surface consists of two parts, or sheets, generated by the two principal centres of curvature. The caustic is simply the surface of centres of the wave surface, and when the surface is one of revolution one of the sheets of the surface of centres degenerates to the axis of revolution.

Now if the whole figure be revolved round OA through a very small angle, PQ will describe a little element of the refracting surface, MN an element of the refracted wave surface, and $F_1$ a short line on the caustic surface perpendicular to the plane of the paper. The element of the surface formed by the revolution of PQ round OA receives a cone of light from O. This cone gives rise to a refracted cone which falls upon the element of the wave surface generated by MN, and the rays of this latter cone produced backwards pass through a short line on the caustic surface at $F_1$.

This line is an element of the circle described by $F_1$ in revolving round the axis OA. It is consequently perpendicular to the plane of the paper—that is, to the primary plane of the refracted cone. (The plane containing the axis of the cone and the axis of the surface is termed the *primary plane* of the cone, while a plane perpendicular to this through the axis of the cone is called the *secondary plane*.) The cross section of the refracted cone at $F_1$ is consequently termed the *primary focal line* of the cone. Now in the case of a surface of revolution it is clear that every normal to the surface meets the axis of revolution, and the point in which any normal $NF_2$ meets the axis remains fixed while the curve BMN revolves. The point $F_2$ is therefore the vertex of a cone of rays passing through the circle described by N about the axis OAB. Similarly every other point on the axis is the vertex of a right circular cone of rays cutting the wave surface orthogonally. Hence the wave normals at two adjacent points M and N meet the axis of revolution at two adjacent points, embracing bebetween them an element of the axis which remains the same while the curve BMN revolves round the axis. The elementary refracted cone of rays previously considered consequently passes through a second line at $F_2$—namely, an element of the axis of revolution. This is the *secondary focal line* of the refracted cone. It lies in the plane of the paper, and is consequently perpendicular to the primary focal line.[1] The cross section of the refracted cone at $F_2$ by a plane perpendicular to its axis is not a line, but in general a figure-of-eight-shaped curve.

Focal lines.

As the section of the refracted cone degenerates to a line at $F_1$ and again to a line at $F_2$, and as these lines are at right angles, it follows that any cross section between $F_1$ and $F_2$ will be an oval curve having diameters in and perpendicular to the primary plane, which reduce to

[1] In the general case when the wave surface is not one of revolution the normals to the surface (*i.e.* the rays) intersect each other only when taken along the lines of curvature. These lines form two orthogonal systems on the surface, and the corresponding intersections of rays form a surface of two sheets (the caustic surface or surface of centres). The primary focal line $F_1$ is an element of a line on one sheet and $F_2$ is an element of a line on the other.

zero as the cross section moves to $F_1$ or $F_2$ respectively. Consequently at some place between $F_1$ and $F_2$ the diameters of the cross section will be equal. The section will here be a circle, or approximately such. This section is called the *circle of least confusion*. Circle of least confusion.

The existence of the focal lines and the circle of least confusion can be easily ascertained by reflecting obliquely from a concave mirror a cone of light diverging from a bright point, and receiving the reflected cone on a white screen or sheet of paper. For one position of the screen a well-defined line perpendicular to the primary plane is depicted. This is the primary focal line. As the screen is moved away farther from the surface the line broadens out into an oval spot, which in one position of the screen is very nearly a circle—the circle of least confusion. On moving the screen farther away the oval narrows in perpendicular to the primary plane till it again becomes a line lying in the primary plane. This is the secondary focal line.

When the incidence of a small cone is direct the focal lines and circle of least confusion all coincide with the geometrical focus C (Fig. 60). This is the point where the caustic meets the axis of revolution. Cusp.
It is a double point or cusp on the caustic curve.

Caustics by reflection may be easily shown by allowing the light of the sun or lamp to fall upon a narrow riband of polished steel such as a watch spring. Placed on a sheet of paper and bent into any desired curve, the spring shows a well-defined bright caustic on the paper, the part within being brighter than that without the curve. By varying the form of the spring a variety of caustics with cusps, contrary flexures, and other singularities may be exhibited. The plane of the sheet of paper should pass nearly through the sun or source of light. The bright curve seen upon the surface of a cup of tea is a familiar example of the caustic of a circle.

## *Examples*

1. Light diverging from a point is refracted at a plane surface, prove that the caustic curve is the evolute of a conic section.

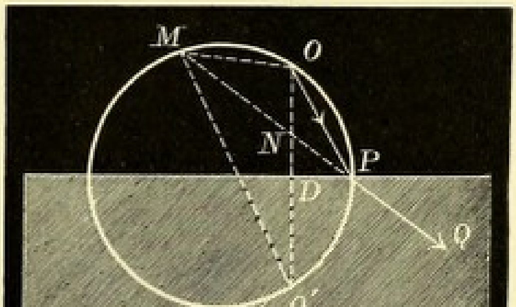


Fig. 61.

[Let O (Fig. 61) be the luminous point, PD the reflecting surface, OP any incident ray. Draw OD perpendicular to the surface, and produce it till $DO' = DO$. Describe a circle about $OPO'$ and produce the refracted ray PQ backwards to meet the circle at M. Join OM and O'M. Then since D is the middle point of OO', the arc $OP = O'P$ or the angle $PMO = PMO'$, that is, the refracted ray bisects the angle M. Now $PMO = POO' = i$, the angle of incidence, and $PND = r$, the angle of refraction. Hence

$$\mu=\frac{\sin i}{\sin r}=\frac{ON}{OM}=\frac{O'N}{O'M},$$

since MN bisects the angle M. Consequently

$$\mu=\frac{ON+O'N}{OM+O'M}, \quad \text{or } OM+O'M=\frac{OO'}{\mu}=\text{constant}.$$

The locus of M is therefore an ellipse having O and O′ for foci, and the refracted ray PQ is a normal to the ellipse at M. The caustic is therefore the evolute of this conic. The major diameter $2a$ and the eccentricity $e$ of the conic are given by

$$2a=OO'/\mu, \quad \text{and } e=\mu.$$

If the refraction takes place from the rarer to the denser medium, the refracted ray bisects the angle M externally. M lies between O and P, and MO′ − MO is constant. The locus of M is a hyperbola, with O and O′ for foci, major diameter OO′/$\mu$, and eccentricity $\mu$.

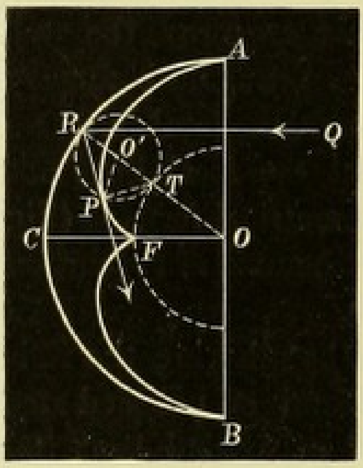


Fig. 62.

The refracted wave front is got by finding a series of points Q such that OP + $\mu$PQ = const. The locus of the points Q will be the refracted wave front cutting the rays PQ normally. The caustic surface is formed by the revolution of the foregoing evolute round OO′.

The reflected waves in this case are spheres concentric with O′, so that the reflected rays all appear to come from O′.]

2. Parallel rays are reflected from a spherical mirror, find the caustic.

[Let ACB (Fig. 62) be a section of the mirror, QR the direction of the incident light, and RP a reflected ray. With centre O (the centre of the mirror) describe a circle of radius OF = ½OC. Join RO and upon RT as diameter describe a circle, centre O′. Then O′T = ½OT. Let the reflected ray cut this circle at P. Now the angle PO′T = 2PRT = 2QRO = 2TOF. Therefore in length the arc TF = arc TP.

If therefore the circle RPT rolls on the circle TF the point P will trace out the epicycloid APFB, the cusp of which is at F, and since RPT is right, and PT is the normal to the path of P, it follows that RP is a tangent to the path. Hence the reflected rays envelop an epicycloid, the cusp of which is at the principal focus of the mirror.]

3. Light diverging from a point on the circumference of a circle is reflected from the concave arc, find the caustic.

[The caustic is an epicycloid formed by the revolution of a circle on an equal circle.]

4. Light diverges from a luminous point situated outside a sphere, find the caustic by refraction.

[A surface of revolution generated by the evolute of a Cartesian oval whose foci are the luminous point and its inverse with respect to the sphere.]

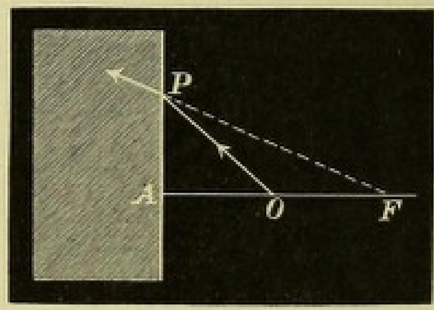


Fig. 63.

5. Find the geometrical focus of a pencil of light after direct refraction at a plane surface.

[Let the light diverge from O (Fig. 63), and let OP be any ray nearly parallel to OA, the perpendicular from O on the surface. Then PA is small, and OP = OA nearly. If the refracted ray produced backwards meets OA at F, then

$$\frac{FA}{OA}=\frac{PF}{PO}=\frac{\sin i}{\sin r}=\mu,$$

so that if OA and FA be denoted by $\rho_1$ and $\rho_2$ respectively, we have

$$\rho_2=\mu\rho_1.$$

The focus is therefore farther away from or nearer to the surface than the point O according as $\mu$ is > or < 1, that is, according as the light passes from the less to the more refracting medium, or *vice versâ*.

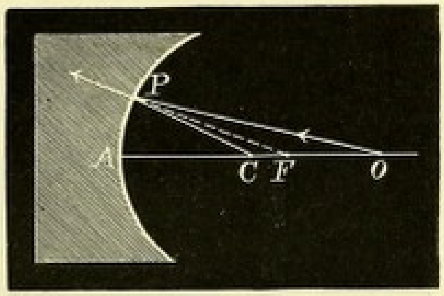


Fig. 64.

To an eye situated in air objects under water appear nearer the surface than they really are, while to an eye placed under water objects in the air appear farther away.]

6. Find the geometrical focus of a pencil of rays after direct refraction at a spherical surface.

[Let C (Fig. 64) be the centre of the sphere and P any point in it, so that OP is nearly parallel to OC. Then if the refracted ray meets OC at F, we have CPO $=i$ and CPF $=r$, and

$$\frac{PF}{PO}=\frac{\sin O}{\sin F}=\frac{\mu CF}{CO},$$

or

$$\frac{\rho_2}{\rho_1}=\frac{\mu(\rho_2-\rho)}{(\rho_1-\rho)}, \quad \therefore \frac{\mu}{\rho_2}-\frac{1}{\rho_1}=\frac{\mu-1}{\rho}.$$

This formula coincides with that which determines the position of the focus by reflection at a spherical mirror when we put $\mu=-1$, a general substitution which converts formulæ of refraction into the corresponding formulæ of reflection.]

7. Light diverges from a point $F_1$; to find the surface which will refract it accurately to another given point $F_2$.

[If $\rho_1$ and $\rho_2$ be the distances of $F_1$ and $F_2$ from any point P of the surface, then the time along $\rho_1$ added to the time along $\rho_2$ must be constant, or

$$\frac{\rho_1}{v_1}+\frac{\rho_2}{v_2}=\text{constant}.$$

That is,

$$\rho_1+\mu\rho_2=\kappa.$$

This is the equation of a Cartesian oval of which $F_1$ and $F_2$ are the foci.

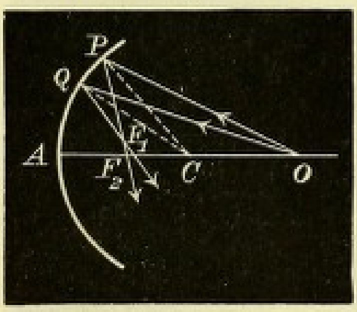


Fig. 65.

If the light be parallel, instead of diverging from a point $F_1$, this surface becomes a spheroid of eccentricity $1/\mu$, and having $F_2$ for a focus.

These surfaces were originally studied by Newton and Descartes, and are termed *aplanatic surfaces* (see Newton's *Principia*, book i. § 14, prop. 97).]

8. A small oblique pencil is reflected at a spherical surface to determine the primary and secondary foci.

[Let C (Fig. 65) be the centre of the surface, O the origin of light, OP and OQ incident rays, and $PF_1$ and $QF_1$ the corresponding reflected rays. $F_1$ the primary and $F_2$ the secondary focus. If OP $=\rho$, $PF_1=\rho_1$, $PF_2=\rho_2$, CP $=$ R, PCA $=\theta$. Then since CP and CQ bisect the angles at P and Q respectively, we have

$$2PCQ=POQ+PF_1Q,$$

but the perpendicular from P on OQ is equal to $\rho$ . POQ, and also to PQ $\cos i$. Therefore $\text{POQ}=\text{PQ}\cos i/\rho$. Similarly $\text{PF}_1\text{Q}=\text{PQ}\cos i/\rho_1$, while $\text{PCQ}=\text{PQ}/\text{R}$. Therefore

$$\frac{1}{\rho_1}+\frac{1}{\rho}=\frac{2}{\text{R}\cos i}.$$

Further to determine $\rho_2$ we have the triangle $\text{OPF}_2=\text{OPC}+\text{CPF}_2$, or $\rho\text{R}\sin i+\text{R}\rho_2\sin i=\rho\rho_2\sin 2i$. Hence

$$\frac{1}{\rho_2}+\frac{1}{\rho}=\frac{2\cos i}{\text{R}}.\ ]$$

9. A small pencil is incident obliquely on a plane refracting surface, find the primary and secondary foci.

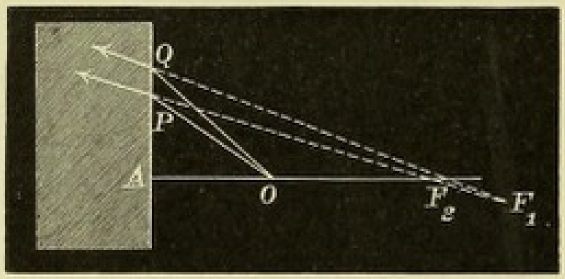



Fig. 66.

[Let OP and OQ (Fig. 66) be two incident rays, $\text{PF}_1$ and $\text{QF}_1$ the refracted. Then if $\text{PF}_1=\rho_1$, $\text{PF}_2=\rho_2$, $\text{OP}=\rho$, we have

$$\frac{\sin i}{\sin r}=\frac{\text{PF}_2}{\text{PO}}=\frac{\rho_2}{\rho},\quad \text{or } \rho_2=\mu\rho,$$

which may be written in the form

$$\frac{\mu}{\rho_2}-\frac{1}{\rho}=0.$$

Again, if we denote the angle $\text{PF}_1\text{Q}$ by $dr$ and POQ by $di$, we have from the triangles $\text{PF}_1\text{Q}$ and POQ

$$\frac{\text{PQ}}{\rho_1}=\frac{dr}{\cos r},\quad \text{and } \frac{\text{PQ}}{\rho}=\frac{di}{\cos i}.$$

But $\sin i=\mu\sin r$, therefore $\cos i\,di=\mu\cos r\,dr$, consequently

$$\frac{\mu\cos^2 r}{\rho_1}-\frac{\cos^2 i}{\rho}=0,$$

or

$$\rho_1=\mu\rho\frac{\cos^2 r}{\cos^2 i}.\ ]$$

10. A small pencil of light is refracted obliquely at a spherical surface of radius R, find the foci.

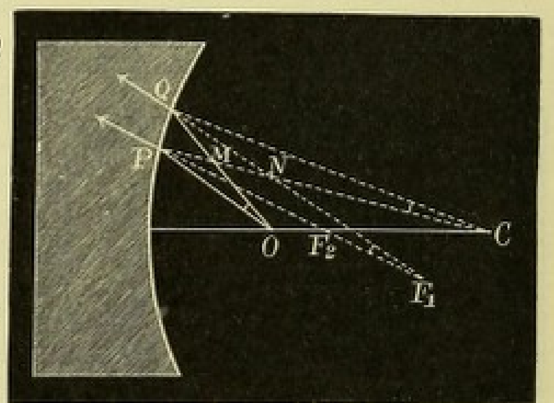



Fig. 67.

[Denote the angles subtended by PQ at O, $\text{F}_1$, and C, by $\alpha$, $\beta$, $\gamma$, respectively. Then, since these angles are very small, we have (Fig. 67)

$$\alpha=\frac{\text{PQ}\cos i}{\rho},\quad \beta=\frac{\text{PQ}\cos r}{\rho_1},\quad \gamma=\frac{\text{PQ}}{\text{R}}.$$

But since the triangles PMO and QMC have equal vertical angles, the sum of the base angles of one is equal to the sum of the base angles of the other, or

$$\alpha+i=\gamma+i+di,\quad \therefore\ di=\alpha-\gamma.$$

Similarly from the triangles $\text{PNF}_1$ and QNC we have

$$\beta+r=\gamma+r+dr,\quad \therefore\ dr=\beta-\gamma.$$

Substituting for $\alpha$, $\beta$, $\gamma$, we find

$$di = PQ\left(\frac{\cos i}{\rho} - \frac{1}{R}\right), \quad dr = PQ\left(\frac{\cos r}{\rho_1} - \frac{1}{R}\right).$$

But since $\sin i = \mu \sin r$, we have

$$\cos i\, di = \mu \cos r\, dr.$$

Therefore

$$\cos i\left(\frac{\cos i}{\rho} - \frac{1}{R}\right) = \mu \cos r\left(\frac{\cos r}{\rho_1} - \frac{1}{R}\right),$$

or

$$\frac{\mu \cos^2 r}{\rho_1} - \frac{\cos^2 i}{\rho} = \frac{\mu \cos r - \cos i}{R}.$$

Again, since the area of the triangle OPC is equal to the sum of the areas of $OPF_2$ and $CPF_2$, we have

$$\rho R \sin i = \rho\rho_2 \sin(i - r) + R\rho_2 \sin r,$$

or dividing across by $\rho\rho_2 R \sin r$, we find

$$\frac{\mu}{\rho_2} = \frac{1}{R}(\mu \cos r - \cos i) + \frac{1}{\rho},$$

or

$$\frac{\mu}{\rho_2} - \frac{1}{\rho} = \frac{\mu \cos r - \cos i}{R}.$$

Hence both results are incorporated in the equations

$$\frac{\mu \cos^2 r}{\rho_1} - \frac{\cos^2 i}{\rho} = \frac{\mu \cos r - \cos i}{R} = \frac{\mu}{\rho_2} - \frac{1}{\rho}.]$$

11. Determine the foci of a small pencil refracted obliquely through a prism in a principal plane.

[Using 9 and neglecting the thickness of the prism we find

$$\rho_1 = \frac{\cos^2 r \cos^2 i'}{\cos^2 i \cos^2 r'}\rho,$$

$$\rho_2 = \rho.$$

Hence if the prism be in the position of minimum deviation (Art. 78) $i = i'$, and $r = r'$, and we have

$$\rho_1 = \rho_2 = \rho,$$

or the foci coincide. The refracted pencil therefore diverges from a point at the same distance ($\rho$) as the origin from the edge of the prism. This result is of importance in the study of the spectrum.

If the refracting angle of the prism be very small, the angles $i$, $i'$, $r$, $r'$ are very nearly equal, and we have again $\rho_1 = \rho_2 = \rho$.]

12. Prove that for a prism of angle A

$$\frac{\sin \frac{1}{2}(A + \delta)}{\sin \frac{1}{2}A} = \mu \frac{\cos \frac{1}{2}(r - r')}{\cos \frac{1}{2}(i - i')}.$$

[We have $\delta = i + i' - A$, and $r + r' = A$, therefore

$$\frac{\sin \frac{1}{2}(A + \delta)}{\sin \frac{1}{2}A} = \frac{\sin \frac{1}{2}(i + i')}{\sin \frac{1}{2}(r + r')}.$$

But $\sin i = \mu \sin r$, and $\sin i' = \mu \sin r'$, therefore by addition we have

$$\sin \tfrac{1}{2}(i+i') \cos \tfrac{1}{2}(i-i') = \mu \sin \tfrac{1}{2}(r+r') \cos \tfrac{1}{2}(r-r'),$$

and consequently

$$\frac{\sin \frac{1}{2}(i+i')}{\sin \frac{1}{2}(r+r')} = \mu \frac{\cos \frac{1}{2}(r-r')}{\cos \frac{1}{2}(i-i')}, \text{ etc.}$$

The least value of $\cos \frac{1}{2}(r-r')/\cos \frac{1}{2}(i-i')$ is unity, and when this happens $i=i'$ and $r=r'$, or the deviation is a minimum. For if $i>i'$ we have $i-r>i'-r'$, or $i-i'>r-r'$. Hence $\cos \frac{1}{2}(i-i')$ is less than $\cos \frac{1}{2}(r-r')$.]

13. Prove that if the refractive index of a prism be changed by an amount $d\mu$, the deviation $\delta$ will be changed by an amount $d\delta$, where

$$d\delta = \frac{\sin A}{\cos r \cos i'} d\mu.$$

[We have $r+r'=A$, $\sin i = \mu \sin r$, $\sin i' = \mu \sin r'$ where A and $i$ are constants, $\delta = i + i' - A$, and consequently $d\delta = di'$.

For the minimum deviation $r = r' = \frac{1}{2}A$, and the above becomes

$$d\delta = \frac{2 \sin \frac{1}{2}A}{\sqrt{1-\mu^2 \sin^2 \frac{1}{2}A}} d\mu,$$

consequently the angle of dispersion of any two colours increases with the angle of the prism.]

14. If a ray of light pass through a system of prisms of vertical angles $A_1$, $A_2$, etc., and if $\alpha_{12}$, $\alpha_{23}$, etc., denote the angles between the faces of the consecutive prisms, prove that the deviation is given by the equation

$$\delta = i_1 + i'_n + \Sigma\alpha - \Sigma A$$

where $i_1$ is the angle of incidence on the first prism, and $i'_n$ the angle of emergence from the last.

[We have $\delta = \Sigma(i + i' - A)$, and also the relations $i'_1 + i_2 = \alpha_{12}$, $i'_2 + i_3 = \alpha_{23}$, etc., consequently $\Sigma(i+i') - i_1 - i'_n = \Sigma\alpha$, therefore, etc.]

15. Determine when the deviation produced by a given system of prisms will be a minimum.

[By Example 14 the deviation will be a minimum when

$$di_1 + di'_n = 0.$$

Now for each prism $dr + dr' = 0$, $\cos i\,di = \mu \cos r\,dr$, $\cos i'\,di' = \mu \cos r'\,dr'$, and therefore

$$\frac{\cos i}{\cos r} di + \frac{\cos i'}{\cos r'} di' = 0,$$

or, writing $\frac{\cos r \cos i'}{\cos i \cos r'} = f$, we have the series of relations $di_1 + f_1 di'_1 = 0$, $di_2 + f_2 di'_2 = 0$, etc., which with the relations $di'_1 + di_2 = d\alpha_{12} = 0$, etc., become

$$di_1 = f_1 di_2, \quad di_2 = f_2 di_3, \text{ etc.}$$

Hence by multiplication and the relation $di_1 + di'_n = 0$ we have

$$1 = f_1 f_2 f_3 \; . \; . \; . \; f_n.$$

By Example 11 it appears that the primary focal distance is given by the equation $\rho_1 = f_1{}^2\rho$, hence for the system of prisms we will have

$$\rho_1 = (f_1 f_2 f_3 \ . \ . \ . \ f_n)^2 \rho,$$

which in the case of minimum deviation becomes

$$\rho_1 = \rho,$$

or the system in this case is still aplanatic.]

16. When rays issuing from a luminous origin are reflected at a given surface the two sheets of the caustic surface have a finite number of points in common. The intensity of the reflected light is greatest at these points, and each of them is the conjugate focus of an ellipsoid of revolution described with the source of light as one focus and having contact of the second order with the mirror (Hamilton, *loc. cit.*).

17. Along a given ray the intensity varies inversely as the product of the distances from the two foci of the pencil (*ibid.*).

## NEWTON'S EXPERIMENTS

"Lights which differ in colour, differ also in degrees of refrangibility" (Newton, *Opticks*, book i. prop. i. theorem 1).

*Exper.* 1.—I took a black oblong stiff paper terminated by parallel sides, and with a perpendicular right line drawn across from one side to the other, distinguished it into two equal parts. One of these parts I painted with a red colour and the other with a blue. The paper was very black and the colours intense and thickly laid on, that the phenomenon might be more conspicuous. This paper I viewed through a prism of solid glass, whose two sides through which the light passed to the eye were plane and well polished, and contained an angle of about sixty degrees, which angle I call the refracting angle of the prism. And whilst I viewed it, I held it and the prism before a window in such manner that the sides of the paper were parallel to the prism, and both those sides and the prism were parallel to the horizon, and the cross line was also parallel to it; and the light which fell from the window upon the paper made an angle with the paper, equal to that angle which was made with the same paper by the light reflected from it to the eye. Beyond the prism was the wall of the chamber under the window covered over with black cloth, and the cloth was involved in darkness that no light might be reflected from thence, which in passing by the edges of the paper might mingle itself with the light of the paper and obscure the phenomenon thereof. These things being thus ordered, I found that if the refracting angle of the prism be turned upwards, so that the paper may seem to be lifted upwards by the refraction, its blue half will be lifted higher by the refraction than its red half. But if the refracting angle be turned downwards, so that the paper may seem to be carried lower by the refraction, its blue half will be carried something lower thereby than its red half. Wherefore in both cases the light which comes from the blue half of the paper through the prism to the eye does in like circumstances suffer a greater refraction than the light which comes from the red half, and by consequence is more refrangible.

*Exper.* 2.—About the aforesaid paper, whose two halves were painted over with red and blue . . . I lapped several times a slender thread of very black silk in such a manner that the several parts of the thread might appear upon the colours like so many black lines drawn over them, or like long and slender dark shadows cast upon them. I might have drawn black lines with a pen, but the threads were smaller and better defined. . . . I placed a candle to illuminate the paper strongly, for the experiment was tried in the night. . . . Then at a distance of six feet and one or two inches from the paper upon the floor I erected a glass lens four inches and a quarter broad, which might collect the rays coming from the several points of the paper, and make them converge towards so many other points at the same distance of six feet and one or two inches on the other side of the lens, and so form the image of the coloured paper upon a white paper placed there. . . . The aforesaid white paper . . . I moved sometimes towards the lens and sometimes from it, to find the places where the images of the blue and red parts of the coloured paper appeared most distinct. Those places I easily knew by the images of the black lines which I had made by winding the silk about the paper. For the images of those fine and slender lines (which by reason of their blackness were like shadows on the colours) were confused and scarce visible, unless when the colours on either side of each line were terminated most distinctly. Noting, therefore, as diligently as I could, the places where the red and blue images of the coloured paper appeared most distinct, I found that where the red half of the paper appeared distinct, the blue half appeared confused, so that the black lines drawn upon it could scarce be seen; and, on the contrary, where the blue half appeared most distinct the red half appeared confused, so that the black lines drawn upon it were scarce

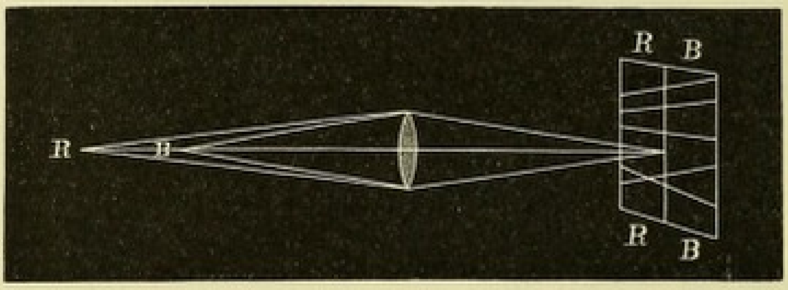


Fig. 68.

Dispersion of foci.

visible. . . . The distance of the white paper from the lens when the image of the red half of the coloured paper appeared most distinct being greater by an inch and a half than the distance of the same white paper from the lens when the image of the blue half appeared most distinct. In like incidences, therefore, of the blue and red upon the lens, the blue was refracted more by the lens than the red, so as to converge sooner by an inch and a half, and therefore is more refrangible.

"The light of the sun consists of rays differently refrangible" (Newton, *Opticks*, book i. prop. ii. theorem 2).

In a very dark chamber, at a round hole about one-third part of an inch broad, made in the shut of a window, I placed a glass prism, whereby the beam of the sun's light, which came in at that hole, might be refracted upwards towards the opposite wall of the chamber, and there form a coloured image of the sun. The axis of the prism (that is, the line passing through the middle of the prism from one end of it to the other end parallel to the edge of the refracting angle)

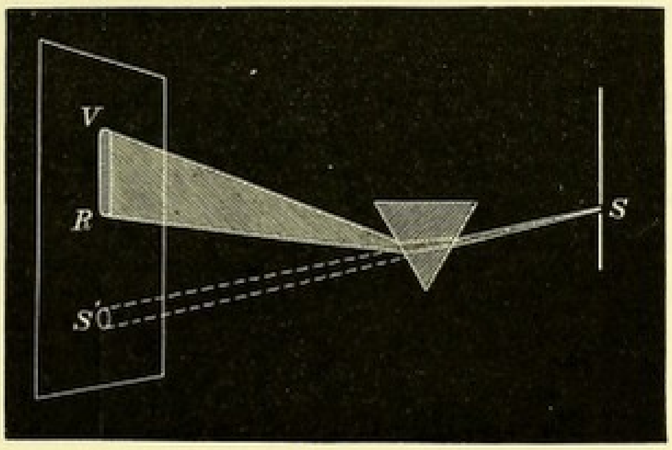


Fig. 69.

was in this and the following experiments perpendicular to the incident rays. About this axis I turned the prism slowly, and saw the refracted light on the wall, or coloured image of the sun, first to descend and then to ascend. Between the descent and ascent when the image seemed stationary, I stopped the prism and fixed it in that posture that it should be moved no more. For in that posture the refractions of the light at the two sides of the refracting angle—that is, at the entrance of the rays into the prism and at their going out of it—were equal to one another. . . . The prism therefore being placed in this posture, I let the refracted light fall perpendicularly upon a sheet of white paper at the opposite wall of the chamber, and observed the figure and dimensions of the solar image formed on the paper by that light. This image was oblong, and not oval, but terminated with two rectilinear and parallel sides, and two semicircular ends. On its sides it was bounded pretty distinctly, but on its ends very confusedly and indistinctly, the light thus decaying and vanishing by degrees. The breadth of this image answered to the sun's diameter, and was about $2\frac{1}{8}$ inches

Minimum deviation.

. . . but the length of the image was about $10\frac{1}{4}$ inches, and the length of the rectilinear sides about 8 inches, and the refracting angle of the prism, whereby so great a length was made, was 64 degrees. With a less angle the length of the image was less, the breadth remaining the same. If the prism was turned about its axis that way which made the rays emerge more obliquely out of the second refracting surface of the prism, the image soon became an inch or two longer, or more; and if the prism was turned about the contrary way, so as to make the rays fall more obliquely on the first refracting surface, the image soon became an inch or two shorter. And, therefore, in trying the experiment I was as curious as I could be in placing the prism by the above-mentioned rule exactly in such a posture that the refractions of the rays at their emergence out of the prism might be equal to that at their incidence on it.

Now the different magnitude of the hole in the window-shut and the different thickness of the prism where the rays passed through it, and the different inclinations of the prism to the horizon, made no sensible changes in the length of the image. Neither did the different matter of the prisms make any; for in a vessel made of polished plates of glass, cemented together in the shape of a prism and filled with water, there is a like success of the experiment according to the quantity of the refraction.

This image of the spectrum was coloured, being red at its least refracted end and violet at its most refracted end, and yellow and blue in the intermediate spaces. . . .

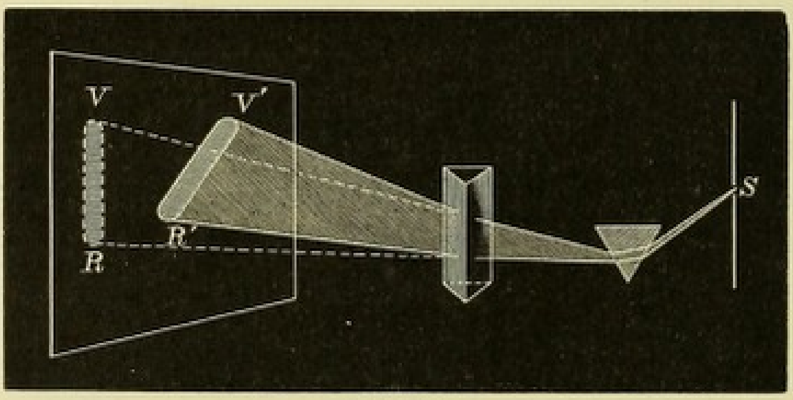


Fig. 70.

I then placed a second prism immediately after the first in a cross position to it, that it might again refract the beam of the sun's light which came to it through the first prism. In the first prism the beam was refracted upwards, and in the second sideways. And I found that by the refraction of the second prism, the breadth of the image was not increased, but its superior part, which in the first prism suffered the greater refraction, and appeared violet and blue, did again in the second prism suffer a greater refraction than its inferior part, which appeared red and yellow, and this without any dilatation of the image in breadth

*Exper.* 6.—In the middle of two thin boards I made round holes a third of an inch in diameter, and in the window-shut a much broader hole being made to let into my darkened chamber a large beam of the sun's light, I placed a prism behind the shut in that beam to refract it towards the opposite wall, and close behind the prism I fixed one of the boards in such a manner that the middle of the refracted light might pass through the hole made in it and the rest be intercepted by the board. Then at a distance of about twelve feet from the first board I fixed the other board in such manner that the middle of the refracted light which came through the hole in the first board and fell upon the opposite wall might pass through the hole in this other board, and the rest being intercepted by the board might paint upon it the coloured spectrum of the sun. And close behind this board I fixed another prism to refract the light which came through

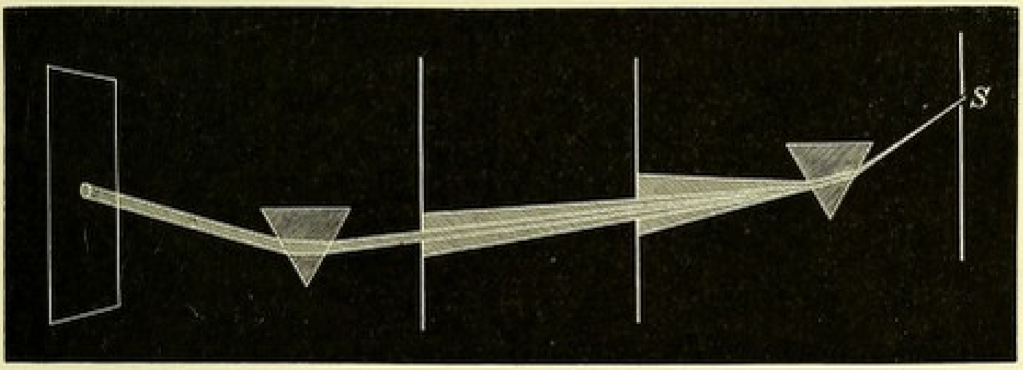



Fig. 71.

the hole. Then I returned speedily to the first prism, and by turning it slowly to and fro about its axis, I caused the image which fell upon the second board to move up and down upon that board that all its parts might successively pass through the hole in that board and fall upon the prism behind it. And in the meantime I noted the places on the opposite wall to which that light after its refraction in the second prism did pass; and by the difference of the places I found that the light which being most refracted in the first prism did go to the blue end of the image, was again more refracted in the second prism than the light which went to the red end of that image, which proves as well the first proposition as the second. And this happened when the axes of the two prisms were parallel or inclined to one another, and to the horizon in any given angles.

"Homogeneal light is refracted regularly without any dilatation splitting or shattering of the rays, and the confused vision of objects seen through refracting bodies by heterogeneal light arises from the different refrangibility of several sorts of rays" (*Opticks*, book i. prop. v. theorem 4).

*Exper.* 12.—In the middle of a black paper I made a round hole about a fifth or sixth part of an inch in diameter. Upon this paper I caused the spectrum of homogeneal light described in the former proposition so to fall that some part of the light might pass through the hole of the paper. This transmitted part of the light I refracted with a prism placed behind the paper, and letting this refracted light fall perpendicularly upon a white paper two or three feet distant from the prism, I found that the spectrum formed on the paper by this light was not oblong, as when it is made (in the third experiment) by refracting the sun's compound light, but was (so far as I could judge by my eye) perfectly circular, the length being no greater than the breadth, which shows that this light is refracted regularly without any dilatation of the rays.

*Note.*—In finding the focal length of a lens by an image cast on a screen, observe that when the screen is too near the edge of the image is red, and when too far the edge is blue; so by this means the screen can be placed very accurately.

## CHAPTER VI

### ON THE DETERMINATION OF REFRACTIVE INDICES

**78. Minimum Deviation.**—Newton, when investigating the solar spectrum, always placed the refracting prism so that the incident and emergent rays were equally inclined to the faces of the prism. In other words, the pencil of light passed symmetrically through the prism so that $i = i'$ and $r = r'$.

In this case we may easily show that the deviation is a minimum. With any centre O (Fig. 72) describe two circles CD and AB with radii proportional to the velocities of light in the two media (air and glass)—that is, the ratio of the radii is equal to the refractive index of the prism. Draw OA parallel to the incident ray, and at A draw CAN parallel to the normal to the first face of the prism. Then $\text{OAN} = i$, and since

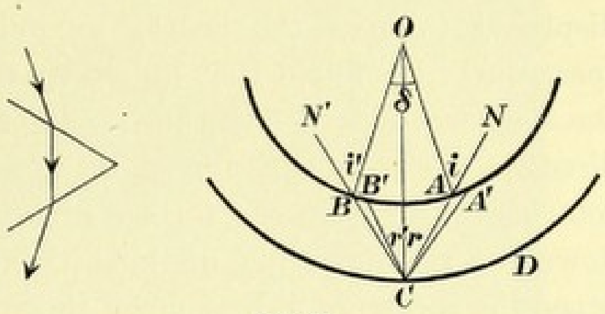


Fig. 72.

$$\frac{\sin i}{\sin r} = \mu = \frac{\text{OC}}{\text{OA}} = \frac{\sin \text{OAN}}{\sin \text{OCA}} = \frac{\sin i}{\sin \text{OCA}},$$

it follows that $\text{OCA} = r$. Consequently OC is the direction of the ray within the prism. If, therefore, CN′ be drawn parallel to the normal to the second face the angle $\text{OCB} = r'$, and since $\text{OC}/\text{OB} = \mu$ it follows as before that $\text{OBN}' = i'$, and therefore OB is parallel to the emergent ray. Hence the angle AOB is equal to the whole deviation, while AOC and BOC are the deviations at the first and second faces respectively, and $\text{ACB} = \text{A}$ the angle of the prism. The angle ACB is therefore constant, and the problem is now reduced to finding when the arc AB is a minimum, for AB measures the angle AOB which is the total deviation.

We shall now show that the arc AB, intercepted on the inner circle by the lines CA and CB inclined at a constant angle A, is least when CA = CB, that is, when the lines are equally inclined to CO. For let CA and CB be this position, and let CA′ and CB′ be a consecutive position, that is, such that the angle ACA′ = BCB′ = $\theta$ where $\theta$ is very small. Then

$$\frac{AA'}{AC} = \frac{\sin\theta}{\sin A'}, \quad \text{and} \ \frac{BB'}{BC} = \frac{\sin\theta}{\sin B'}.$$

Hence if AC = BC we have

$$\frac{AA'}{BB'} = \frac{\sin B'}{\sin A'}.$$

But obviously $\sin B' > \sin B$ and $\sin A' < \sin A$, B′ being an exterior and A′ an interior angle, and B = A, therefore $\sin B' > \sin A'$ or AA′ > BB′. Consequently A′B′ > AB or the arc AB is less than its consecutive value on either side, and is therefore a minimum[1] (see also Ex. 12, p. 115).

In practice the position of minimum deviation is easily determined. Newton, who worked with a small pencil of solar light from an aperture in the shutter, placed the prism edge downwards. Consequently the refracted pencil was bent upwards, and the spot on the screen was displaced towards the ceiling as well as converted into a coloured spectrum. He found that by gently turning the prism round its edge the spectrum moved down the screen, and by rotating it in the opposite direction the spectrum moved up so as to increase the deviation. By rotating it slowly in the former direction so that the spectrum moved down, or the deviation diminished, he found that as he turned the prism the spectrum moved down the screen to a certain position, where it came to a standstill, and then commenced to move upwards again, no matter which way he rotated the prism. This position is that of

[1] Or again, let the circle centre C radius CB′ cut the circle BA again in A″. Then AA′/AA″ = CA′/CA″ because CA bisects A″CA′. Therefore BB′ = AA″ < AA′.

Or thus: since $\delta = i + i' - A$,
we have for $\delta$ a max. or min. $0 = di + di'$,
also, since $r + r' = A$, $0 = dr + dr'$.
Again, $\sin i = \mu \sin r$ and $\sin i' = \mu \sin r'$,
therefore $\cos i\,di = \mu \cos r\,dr$ and $\cos i'\,di' = \mu \cos r'\,dr'$.
Hence by division we have

$$\frac{\cos i}{\cos i'} = \frac{\cos r}{\cos r'}, \quad \therefore \frac{\cos^2 i}{\cos^2 i'} = \frac{\cos^2 r}{\cos^2 r'},$$

or

$$\frac{1 - \mu^2 \sin^2 r}{1 - \mu^2 \sin^2 r'} = \frac{1 - \sin^2 r}{1 - \sin^2 r'} = \frac{\sin^2 r}{\sin^2 r'},$$

the third expression being obtained from the first and second by taking the ratio of the difference of their numerators to the difference of their denominators. Hence $r = r'$, and consequently $i = i'$. It is easily seen that this corresponds to a minimum rather than a maximum value of $\delta$.

the minimum deviation, and is, as Newton observed, that in which the ray passes symmetrically through the prism.

The same principle is used at present, although the method of procedure is different, the spectrum being now viewed through a telescope instead of being projected on a screen.

An instrument fitted to observe a spectrum is called a spectroscope, or a spectrometer if, in addition, it be graduated to measure the deviation of the refracted rays.

**79. The Spectrometer.**—Fig. 73 represents the outline of a spectrometer. The source of light is placed before a narrow slit S in the end of a telescope tube, or if greater illumination be desired the light

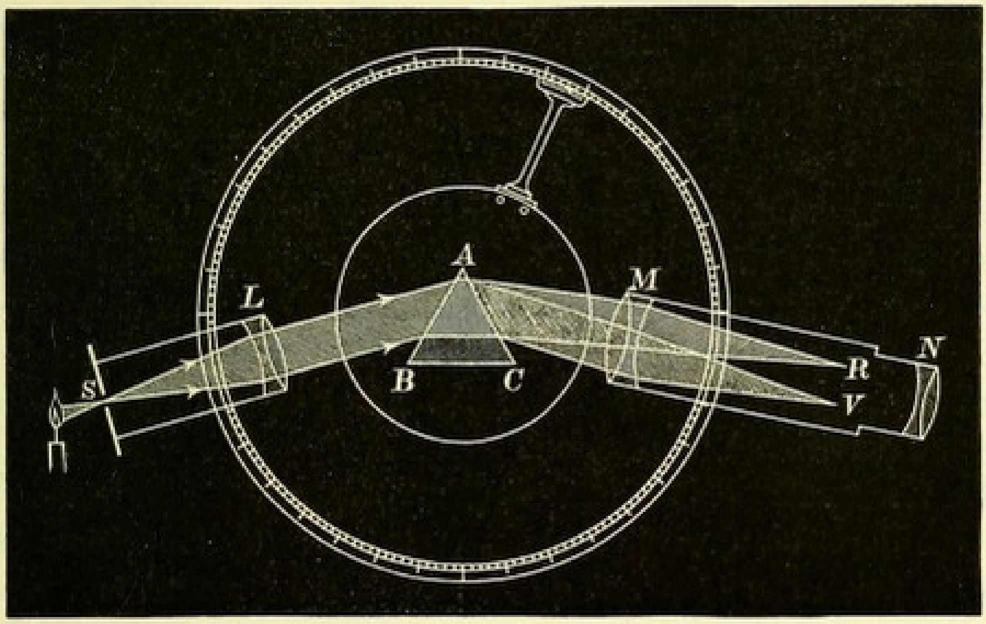



Fig. 73.—The Spectrometer.

of the sun or electric arc is concentrated at S by means of a condensing lens. At the other end of the tube an achromatic lens L is placed and focussed on S, so that the light emerges from L in a parallel beam. This part of the apparatus is called the *collimator*, and its function is to procure a parallel beam of light. The light emerging from the collimator falls upon the face AB of the prism ABC. The prism is placed with its refracting edge (A) vertical on a horizontal table, at the centre of the instrument, which can be levelled by means of three adjusting screws. In passing through the prism the light suffers dispersion, and there emerges a parallel beam of red light and also a parallel beam of violet, with beams of the other colours situated between them. The pencil of red light is brought to a focus R and the violet light to a focus V by the object glass M of the observing telescope. The other colours are focussed at points lying between V and R, so that a real spectrum is depicted in the focal plane of the telescope. This spectrum is then viewed through a suitable eyepiece N.

Since the property of a lens is to concentrate a system of parallel rays to a single focus, we see that the use of the collimator is of great importance in producing a pure spectrum. For since the red light emerges from the prism in a parallel beam, it will form a red image of the slit at R, similarly there will be a violet line at V, and the other coloured images will be spread out between R and V without overlapping if the slit be sufficiently narrow. The dispersion, as the separation of the colours is called, depends on the nature of the prism as well as its angle (see Ex. 13, p. 116), and can be greatly increased by allowing the light to pass through several prisms in succession so that a very long spectrum may be obtained.

The collimator and telescope are attached to the graduated circle, their axes are parallel to its plane, and their directions meet above its centre. The telescope is generally fitted with cross wires and a Ramsden's eyepiece, and turns in a plane parallel to the graduated circle about its centre while the collimator is fixed. The position of the telescope with reference to the circle is determined by a vernier. The table which supports the prism turns round a vertical axis and carries an arm furnished with a vernier, which slides on the graduated circle and determines its position.

**80. Measurement of Minimum Deviation and Angle of Prism.**—Since the deviation depends on the wave frequency (or colour) of the light, the prism can be placed in the position of minimum deviation only for some selected colour at once. To measure the minimum deviation for any ray the telescope is turned to view the slit of the collimator directly, and its reading is then taken on the graduated circle. The prism of the substance of which we wish to determine the refractive index is now fixed on the table of the spectroscope, and the telescope is turned to view the spectrum produced by the prism. The prism is then rotated (by turning the table which supports it) through a small angle, and this will generally either increase or diminish the deviation. Turning it in one direction will increase, and in the opposite will diminish, the deviation. This latter direction being determined, the prism is slowly turned, and the spectrum is followed by the telescope till the spectrum comes to a standstill, and any further turning of the prism in either direction will increase the deviation. The reading of the telescope is now taken, and the difference between this and the former reading is the minimum deviation of the ray under consideration.

To determine A, the angle of the prism, turn the table of the spectroscope so that the edge of the prism (Fig. 74) is towards the collimator, and the light from it falls upon and is reflected from both faces. The telescope M is now turned until the light reflected from the face AB enters it, and an image of the slit is seen by reflection in the field of view coinciding with the cross wires.

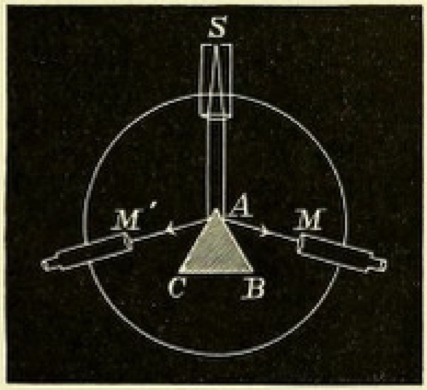


Fig. 74.

Now turn the telescope into the position M′, so that an image of the slit is seen by reflection in the face AC. The difference between the readings of the telescope in these two positions, that is, the angle through which it has been turned, is twice the angle of the prism, for this is only an illustration of the principle that if a plane reflector is turned through any angle the reflected ray is turned through twice that angle (chap. iv. Ex. 1). Or thus: the angle between the reflected ray AM and the incident ray produced is equal to twice the angle between AB and the incident light, so also the angle between AM′ and the incident ray produced is twice the angle between the face AC and the incident light. Therefore the angle between AM and AM′, or the angle through which the telescope has been turned, is equal to twice the angle of the prism.

The prism should be placed on the table, so that the light from the central portion of the collimating lens falls upon its edge, and does not pass by it. At the same time it is to be so adjusted that it is possible to see the slit by reflection from both faces, as described above. These conditions are generally secured by placing the prism with its edge a little in front of the centre of the supporting table.

This method of determining A supposes the faces of the prism to extend right up to its edge, and that near the edge the faces exhibit no curvature. If any doubt exists as to the faces being truly plane at the edge we may employ another method of determining A, which depends upon keeping the telescope fixed and rotating the table. This method may be used when the table is furnished with a vernier arm to determine its position.

Clamp the telescope M in any position, so that its axis is inclined to the axis of the collimator, and turn the table till the image of the slit is seen in the telescope by reflection from the face AB of the prism. Read the vernier and then turn the table until the slit is seen by reflection from the other face AC. Read the table vernier again. The difference between the two readings is easily seen to be $\pi - A$.

Having measured both A and $\delta$, we determine $\mu$ by the formula of the following article. A Bunsen burner with a bead of salt in it gives a yellow sodium light which is very homogeneous. By using it we determine the refractive index for yellow light and avoid complications. To ensure accuracy the determinations of A and $\delta$ should in all cases be repeated several times.

**81. Formula for the Refractive Index.**—The prism being placed in the position of minimum deviation for any particular ray of light we have

$$\delta = 2i - A, \quad \text{or } i = \tfrac{1}{2}(A + \delta),$$

and

$$A = 2r, \quad \text{or } r = \tfrac{1}{2}A,$$

but

$$\sin i = \mu \sin r,$$

therefore

$$\sin \tfrac{1}{2}(A + \delta) = \mu \sin \tfrac{1}{2}A$$

and

$$\mu = \frac{\sin \frac{1}{2}(A + \delta)}{\sin \frac{1}{2}A},$$

a formula which determines $\mu$ when A and $\delta$ have been measured.[1]

*Cor.*—If the prism be very thin, so that A and $\delta$ are very small, the sines of the angles may be replaced by their circular measures, and the above formula becomes

$$\delta = (\mu - 1)A.$$

*Ex.* 1. The minimum deviation $\delta$ varies with the colour or wave length of the light. The change $d\mu$ of the index and the corresponding change $d\delta$ of the minimum deviation are connected by the equation

$$d\mu = \frac{\cos \frac{1}{2}(A + \delta)}{2 \sin \frac{1}{2}A} d\delta.$$

*Ex.* 2. A prism of silicate of lead having a refracting angle of 21° 12′ produced a minimum deviation of 24° 46′ in homogeneous red light; find the refractive index

[L sin $\tfrac{1}{2}(A + \delta)$ = L sin (22° 59′) = 9·59158

L sin $\tfrac{1}{2}$A = L sin (10° 36′) = 9·26470.

Therefore log $\mu$ = 0·32688 and $\mu$ = 2·123].

**82. Liquids.**—The refractive index of a liquid may be determined by enclosing it in a hollow prism the faces of which are plates of

[1] *Method of Descartes.*—The method employed by Descartes to determine the refractive index of a substance is of considerable interest. Transmitting a horizontal pencil of light through an aperture O (Fig. 75) in a vertical screen AB, it fell upon the vertical face of a prism placed against an aperture O′ in the vertical screen A′B′. In this case

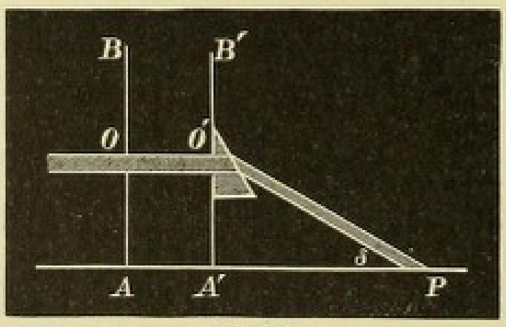


Fig. 75.

$$i = 0, \; r = 0, \; r' = A, \quad \text{and } \delta = i' - A,$$

and therefore

$$\sin (A + \delta) = \mu \sin A.$$

The angle $\delta$ which the refracted pencil makes with the horizon is determined by the equation O′A′ = A′P tan $\delta$ (*Dioptrices*, caput x. p. 140).

parallel glass. If the glass sides of the hollow prism be uniform plates they will not cause any dispersion or deviation of the light. The whole procedure is consequently the same as in the case of transparent solids.[1]

**83. Method of Total Reflection.**—By measuring the critical angle ($\alpha$) of a substance we at once determine its relative index of refraction by the formula

$$\sin \alpha = 1/\mu.$$

The advantage of this method is that it does not require the construction of a prism, and may therefore be applied to a solid substance which it is not convenient to cut. Wollaston[2] applied the method to liquids, and in this application it has been perfected by MM. Terquem and Trannin.[3] Two plates of parallel glass (Fig. 77) cemented together at their edges by a little gum or Canada balsam, so as to enclose a thin layer of air, are immersed in a vessel containing the liquid under consideration. The vessel, which is a square box with parallel glass sides, is first fixed on the table of the spectrometer and the telescope is adjusted to view the collimator slit, the light passing normally through the sides of the box. The prepared glass plates are now placed in the liquid with their plane faces vertical (the table of the spectroscope being horizontal), and so inclined to the light from the collimator that total reflection just occurs. When this happens the Liquids.

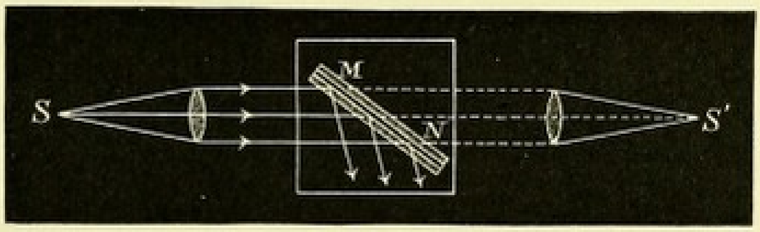


Fig. 77.

[1] *Newton's Method.*—An ingenious method of measuring the refractive index of a liquid was employed by Newton. The liquid was placed in a rectangular vessel with a flat glass bottom M (Fig. 76). This vessel is attached to a frame AB, which moves freely round a horizontal axis O, carried by a vertical support OP. A pencil of solar light falls upon the surface of the liquid, and the carrier AB is turned round O till the pencil emerging from the base of the vessel is parallel to AB. In this position the refracted beam is perpendicular to the base M of the vessel. The angle of refraction is consequently equal to the inclination of AB to the vertical. This is determined by a graduated circle attached to OP. The angle of incidence is the inclination of the incident ray to the vertical, and this may be determined by repeating the experiment when there is no liquid in the vessel.

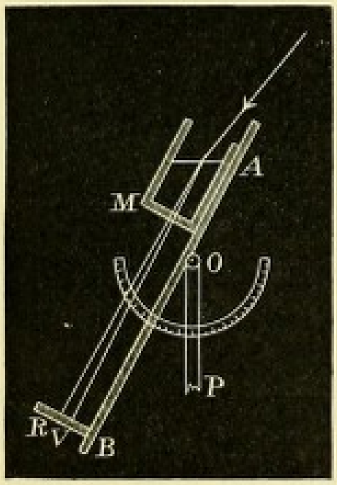


Fig. 76.

[2] Wollaston, *Phil. Trans.* p. 365, 1820.

[3] Terquem and Trannin, *Journal de Physique*, first series, tom. iv. p. 232, 1875.

field of the telescope becomes quite dark. It is clear that in this case the angle of incidence on the glass plates is the angle of total reflection of the liquid and air. By turning the plates round the vertical in the opposite direction, the light reappears in the telescope, and then, on continued rotation of the plates, again vanishes. The angle through which the plates have been turned is twice the critical angle $\alpha$, and

$$\sin \alpha = 1/\mu.$$

Solids. In a similar manner the index of a thin plate of a solid may be determined by immersing it in a liquid of known index $\mu_0$ greater than that of the solid. Here we have

$$\sin \alpha = \mu/\mu_0.$$

This method has been adopted by Kohlrausch for measuring the indices of crystals.[1]

It is important that the light should fall upon the plates containing the air film in a parallel beam, for the whole pencil will then suffer total reflection at the same instant, and the field of the telescope will become dark suddenly. Care should also be taken to place the plates vertically. This being secured, with monochromatic (sodium) light, the disappearance of the image is almost instantaneous, and MM. Terquem and Trannin estimate the experimental error at less than fifteen seconds. When white light is used, the image passes through shades of yellow, orange, and finally the pure red of the extreme spectrum. The accuracy of the method is shown by the following table, where the indices are compared with the determinations of Fraunhofer, and of Gladstone and Dale, the slight differences being attributable to the state of purity of the liquid :—

| | Ray. | Temp. | 2 α. | Index. | |
|---|---|---|---|---|---|
| Water . . . . | C | 18° | 97° 20′ 30″ | 1·3317 | 1·33171 Fraunhofer. |
| ,, . . . . | D | 18° | 97° 9′ 50″ | 1·3336 | 1·33171 ,, |
| Benzine . . | A | 19°·5 | 84° 41′ 20″ | 1·4846 | 1·4860 Gladstone and Dale. |
| Glycerine . . | A | 18° | 85° 55′ 20″ | 1·4672 | 1·4664 ,, ,, |
| Amyllic Alcohol . | A | 18° | 91° 10′ 0″ | 1·4000 | 1·3990 ,, ,, |
| Sulphide of Carbon | A | 20° | 76° 55′ 0″ | 1·6078 | 1·6076 ,, ,, |

When the liquid can be procured in small quantities only, or when it is imperfectly transparent or pasty (such, for example, as the crystalline lens of the eye), Wollaston's original method may be used. A small drop D (Fig. 78) of the liquid is attached to the lower surface

[1] *Wied. Ann.* iv. 1, 1878.

of a right-angled prism ABC, resting on a horizontal table AB furnished with a small hole to receive the drop. The drop is viewed through the face AC of the prism by means of a telescope T pivoted at the centre of a graduated circle, which slides on a vertical rod fixed to the table. The telescope is first placed at such a height that the table, or any mark below the drop, may be seen through it. As the graduated circle is lowered the angle of reflection from the drop is increased, till at last the rays falling on it from the other side of the prism are totally reflected. The telescope is now fixed, and its inclination ($\theta$) to the vertical is observed. In this case NDM is the critical angle $\alpha$, and $NMD = 90 - \alpha$, while $TMO = 90 - \theta$. But the refractive index of the glass is given by

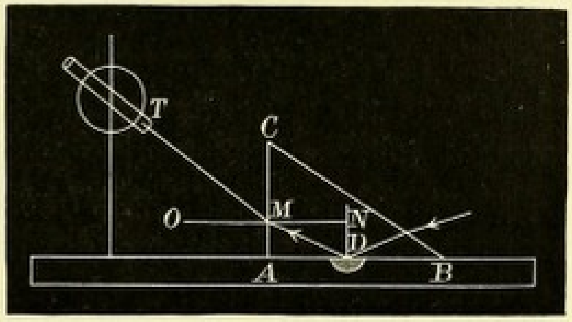



Fig. 78.

$$\mu = \frac{\sin TMO}{\sin NMD} = \frac{\cos\theta}{\cos\alpha}.$$

Hence

$$\cos\alpha = \frac{\cos\theta}{\mu}, \quad \text{and } \sin\alpha = \frac{1}{\mu}\sqrt{\mu^2 - \cos^2\theta}.$$

But if the refractive index of the drop be denoted by $\mu'$, we have

$$\mu' = \mu \sin\alpha = \sqrt{\mu^2 - \cos^2\theta}.$$

An objection to the method is the difficulty of making the angle of the prism exactly right. To avoid this, Malus worked with an acute-angled prism, of which the index and angle were accurately determined.

*Examples*

1. When the angle A of the prism is acute, prove that

$$\mu' = \sin(\theta - A)\cos A + \sin A\sqrt{\mu^2 - \sin^2(\theta - A)}$$

where $\theta$, as before, is the angle the telescope makes with the vertical.

[The angle of incidence of the ray DM on the face AC is $r = A - \alpha$, and the angle of emergence is $i = A - \theta$. But $\sin i = \mu \sin r$, therefore $\sin(A - \theta) = \mu \sin(A - \alpha)$, and

$$\mu' = \mu\sin\alpha = \mu\sin(A - r) = -\mu\sin r\cos A + \mu\sin A\cos r, \text{ etc.}]$$

2. Determine the refractive index of the prism by the same method.

[Since the medium below the prism is air, we have $\mu' = 1$, and

$$\mu = \operatorname{cosec} A\sqrt{1 + \sin^2(\theta - A) - 2\cos A\sin(\theta - A)}.]$$

**84. Gladstone and Dale's Law—Variation of Refractive Index with Density.**—When a substance is compressed, or its temperature varied, the density ($\rho$) alters, and this is accompanied by a corresponding variation in the refractive index ($\mu$). Gladstone[1] and Dale found that these two quantities were related by the equation

$$\frac{\mu - 1}{\rho} = \text{constant}.$$

The physical interpretation of this law is not far to seek if we admit that the refracting substance consists of refracting molecules of constant index distributed through the ether. The quantity $\mu - 1$ represents the excess of the refraction or path retardation due to the presence of the molecules, and will be proportional to their number per unit volume, that is, to the density. Thus if $\mu$ be the index of a plate of the substance and $e$ its thickness, the corresponding path in free space of a ray traversing the plate will be $\mu e$, but if the thickness absolutely occupied by the molecules of the substance be $\epsilon$ and the constant index[2] of a molecule be $m$, then a thickness $e - \epsilon$ of the plate is occupied by ether, and the corresponding free space path for the plate will be $e - \epsilon + m\epsilon$. Hence

$$\mu e = e - \epsilon + m\epsilon,$$

or

$$\mu - 1 = \frac{\epsilon}{e}(m - 1).$$

But $\epsilon/e$ measures the relative volume occupied by the molecules, and is therefore proportional to the density of the body. Consequently

$$\frac{\mu - 1}{\rho} = \text{constant}.$$

Since the density of a solid or liquid can be changed only very slightly, experiments on the law are necessarily confined within extremely narrow limits, and we can only accept it as an approximation to the truth in absence of other support.

In the case of gases, however, the density can be varied considerably, but on the other hand $\mu - 1$ is here a very small quantity, and therefore experiments which sufficiently verify the formula

$$\frac{\mu - 1}{\rho} = \kappa$$

will also verify the formula

$$\frac{\mu^2 - 1}{\rho} = 2\kappa,$$

for $\mu + 1 = 2$ very nearly, if $\mu$ exceeds unity by a small amount.

[1] Gladstone and Dale, *Phil. Trans.* p. 887, 1858, and p. 317, 1863.

[2] The quantity $\epsilon$ is the thickness of the plate of the same face area that the molecules of the plate $e$ would form if arranged with no free spaces between them, and the $m$ is the index of this derived plate.

The emission theory indicates that for all bodies the quantity $\mu^2 - 1$ should be proportional to the density, and consequently it has been sought to establish the latter formula. Thus, if $v'$ be the velocity in a given substance and $v$ the velocity in a vacuum, the work done on a luminous corpuscle in passing into the substance from vacuum will be proportional to the density of the substance, but the work done on the corpuscle is proportional to the change in the square of its velocity or *vis viva*, therefore

$$v'^2 - v^2 = \kappa\rho.$$

That is,

$$\mu^2 - 1 = \kappa\rho/v^2.$$

This shows that the index is independent of the angle of incidence, and that $\mu^2 - 1$ is directly proportional to the density of the substance.

From his experiments on the variation of the index of water under compression, M. Jamin concluded that $(\mu^2 - 1)/\rho$ was constant, but M. Mascart found that the results of M. Jamin were more accurately represented by the formula of Gladstone and Dale, and this conclusion has been verified by Quincke.[1]

*Effect of Variation of Temperature.*—If the temperature varies the molecular index may also vary. If this be so it follows from the equation

$$\mu - 1 = v(m - 1),$$

where $v$ is the volume of the molecules in a unit volume of the substance, by differentiating logarithmically with regard to the temperature $\theta$, that

$$\frac{1}{\mu - 1}\frac{d\mu}{d\theta} = \frac{1}{v}\frac{dv}{d\theta} + \frac{1}{m - 1}\frac{dm}{d\theta}.$$

But the density of the medium is proportional to $v$. Hence

$$\frac{1}{\mu - 1}\frac{d\mu}{d\theta} - \frac{1}{\rho}\frac{d\rho}{d\theta} = \frac{1}{m - 1}\frac{dm}{d\theta},$$

or, what is more commodious for calculation, if V be the volume at $\theta$ of a unit mass of the substance, then $V = 1/\rho$, and

$$\frac{1}{\mu - 1}\frac{d\mu}{d\theta} + \frac{1}{V}\frac{dV}{d\theta} = \frac{1}{m - 1}\frac{dm}{d\theta}.$$

---

[1] Ketteler proposes the empirical formula

$$\frac{\mu^2 - 1}{\rho}(1 - \alpha\rho - \beta\rho^2 - \gamma\rho^3 \ . \ . \ . \ ) = \text{constant}.$$

The ratio $(\mu^2 - 1)/(\mu'^2 - 1)$ for the different colours he finds to be sensibly constant so that $\alpha$, $\beta$, $\gamma$, etc. do not depend on the wave length. In the case of water below 8° they appear, however, to depend on the temperature (*Wied. Ann.* xxx. p. 299, 1887 ; *Journal de Physique*, second series, tom. vii. 1888).

Gladstone's law, which assumes $m$ to be constant, requires the second member to vanish. M. Dufet[1] found it always negative and very nearly constant for liquids, but greater in absolute value than indicated by theory when $m$ is a variable. In the case of solids the sign was positive and the magnitude less than that deduced from theory, whereas in the case of solutions when concentrated they behaved as solids, but when dilute they behaved as liquids.

Or in all liquids the molecular index decreases as the temperature rises, and in all solids the molecular index increases with the temperature while the quantity $\frac{1}{m-1}\frac{dm}{d\theta}$ remains very nearly constant.

The effect of temperature on the refractive index is shown in the following table after Gladstone and Dale. The indices are for the D line, except that of phosphorus, which is for the C line :—

| Temperature. | Sulphide of Carbon. | Water. | Ether. | Absolute Alcohol. | Methyl Alcohol. | Phosphorus (Liquid). |
|---|---|---|---|---|---|---|
| 0° | 1·6442 | 1·3330 | ... | ... | ... | .. |
| 10° | 1·6346 | 1·3327 | 1·3592 | 1·3658 | 1·3379 | ... |
| 20° | 1·6261 | 1·3320 | 1·3545 | 1·3615 | ... | ... |
| 30° | 1·6182 | 1·3309 | 1·3495 | 1·3578 | ... | 2·0741 |
| 40° | 1·6103 | 1·3297 | ... | 1·3536 | 1·3297 | 2·0677 |
| 50° | ... | 1·3280 | ... | 1·3491 | ... | 2·0603 |
| 60° | ... | 1·3259 | ... | 1·3437 | ... | 2·0515 |

**85. Indices of Mixtures and Solutions.**—The law of Gladstone may be applied to calculate the refractive index of a mixture of two substances which have no chemical action on each other.

Let $V_1$ and $V_2$ be the volumes, $\rho_1$ and $\rho_2$ the densities, $\mu_1$ and $\mu_2$ the indices of the components of the mixture, and let $V$, $\rho$, $\mu$ be the corresponding quantities for the mixture, we have

$$\rho V = V_1\rho_1 + V_2\rho_2,$$

and since each of the components occupies a volume V in the mixture we can say that $V_1\rho_1/V$ and $V_2\rho_2/V$ are the densities of the components in the mixture. Consequently by Gladstone's law

$$\frac{\mu_1-1}{\rho_1}=\frac{\mu'_1-1}{\rho'_1}=\frac{(\mu'_1-1)}{\rho_1}\frac{V}{V_1}$$

and

$$\frac{\mu_2-1}{\rho_2}=\frac{\mu'_2-1}{\rho'_2}=\frac{(\mu'_2-1)}{\rho_2}\frac{V}{V_2},$$

[1] "Sur la loi de Gladstone et la Variation de l'Index moléculaire." Par M. H. Dufet, *Journal de Physique*, November 1885. See also September 1885.

where $\mu'_1$ and $\mu'_2$ are the new indices of the components in the mixture according to the law. The index $\mu$ of the mixture will therefore be given by $(\mu - 1) = (\mu'_1 - 1) + (\mu'_2 - 1)$, or by the foregoing equations

$$\mu - 1 = (\mu_1 - 1)\frac{V_1}{V} + (\mu_2 - 1)\frac{V_2}{V},$$

or, finally, if $M_1$ and $M_2$ be the masses of the components,

$$(M_1 + M_2)\frac{\mu - 1}{\rho} = M_1\frac{\mu_1 - 1}{\rho_1} + M_2\frac{\mu_2 - 1}{\rho_2}.$$

Experiment shows that this formula is very approximately true for mixtures of liquids and mixtures of isomorphous crystallised salts, but it ceases to be exact for mixtures of saline solutions.

*Effect of Change of Temperature.*—If the temperature ($\theta$) varies, we have

$$(M_1 + M_2)\frac{d}{d\theta}\left(\frac{\mu - 1}{\rho}\right) = M_1\frac{d}{d\theta}\left(\frac{\mu_1 - 1}{\rho_1}\right) + M_2\frac{d}{d\theta}\left(\frac{\mu_2 - 1}{\rho_2}\right).$$

But by the (equation $\mu - 1) = v(m - 1)$ it follows that

$$\frac{d}{d\theta}\left(\frac{\mu - 1}{\rho}\right) = \frac{v}{\rho}\frac{dm}{d\theta} = \frac{(\mu - 1)}{\rho}\cdot\frac{1}{m - 1}\frac{dm}{d\theta}.$$

Hence

$$(M_1 + M_2)\frac{\mu - 1}{\rho}\cdot\frac{1}{m - 1}\frac{dm}{d\theta} = M_1\frac{\mu_1 - 1}{\rho_1}\cdot\frac{1}{m_1 - 1}\frac{dm_1}{d\theta} + M_2\frac{\mu_2 - 1}{\rho_2}\cdot\frac{1}{m_2 - 1}\frac{dm_2}{d\theta}.$$

**86. Gases.**—The index of refraction of a gas may be determined, like that of a liquid, by enclosing it in a hollow prism with faces of parallel plate glass. The prism is placed on the table of the spectrometer, and the gas under consideration is admitted through drying tubes, if necessary, its temperature and pressure being noted. The telescope, as usual, is first adjusted to observe the collimator slit directly, and if the plate sides of the prism are of truly parallel glass, the prism, when placed on the table of the spectroscope, will produce no deviation when it contains free air. This being ensured, any future deviation is due to the refractive difference between the gas contained by the prism and the outside air. Any such deviation will in general be very small, so that the table of the instrument (with the prism attached) may be turned through 180°. This will throw the deviation to the opposite side (as the edge of the prism is now turned in the opposite direction), so that the difference between the two readings of the telescope will be $2\delta$, and by means of the ordinary formula we obtain the relative index of the gas.

To obtain its absolute index we require first the index of air at zero and 760 mm. This is determined by first pumping as much as possible of the air out of the prism. Let the pressure of the residual air be $h'$, its index $\mu'$, and the density $\rho'$, then by the law of Gladstone and Dale

$$\frac{\mu'-1}{\rho'}=\frac{\mu_0-1}{\rho_0},$$

or $$\mu'=1+(\mu_0-1)\frac{\rho'}{\rho_0}=1+(\mu_0-1)\frac{h'}{760(1+\alpha\theta)},$$

if $\theta$ be the temperature.

But if the exterior air be at a pressure $h$, density $\rho$, and index $\mu$, we have also

$$\frac{\mu-1}{\rho}=\frac{\mu_0-1}{\rho_0},$$

or $$\mu=1+(\mu_0-1)\frac{\rho}{\rho_0}=1+(\mu_0-1)\frac{h}{760(1+\alpha\theta)}.$$

Experiment gives the relative index of the air in the prism with respect to the exterior air; if this be $m$ we have

$$m=\frac{\mu'}{\mu}=\frac{760(1+\alpha\theta)+(\mu_0-1)h'}{760(1+\alpha\theta)+(\mu_0-1)h}.$$

Hence $$\mu_0=1+\frac{(m-1)760(1+\alpha\theta)}{h'-mh}.$$

To determine the refractive index $\nu_0$ of any gas at zero and 760 (the pressure being $h'$), we have as before, by Gladstone's law,

$$\nu'=1+(\nu_0-1)\frac{h'}{(1+\alpha\theta)760},$$

where $\nu'$ is the index at $h'$ and $\theta$.

If the outside air be at $h$ and $\theta$, we have for it, as above,

$$\mu=1+(\mu_0-1)\frac{h}{(1+\alpha\theta)760}.$$

Consequently if $m_1$ be the relative index of the gas with respect to the air, we have

$$m_1=\frac{\nu'}{\mu}=\frac{(1+\alpha\theta)760+(\nu_0-1)h'}{(1+\alpha\theta)760+(\mu_0-1)h},$$

from which we obtain $\nu_0$.

Dulong[1] worked with a constant deviation. He introduced the different gases successively into a hollow prism, but varied the pressure so that each produced the same deviation. Thus for air

[1] *Annales de Chimie et de Physique*, second series, tom. xxxi. p. 154, 1826.

at $h'$ and $\theta$ in the prism, while the exterior air is at $h$ and $\theta$, we have for the relative index

$$m=\frac{(1+\alpha\theta)760+(\mu_0-1)h'}{(1+\alpha\theta)760+(\mu_0-1)h},$$

while for a gas at $h'_1$ and $\theta_1$, with relative index $m_1$, we have

$$m_1=\frac{(1+\alpha\theta_1)760+(\nu_0-1)h'_1}{(1+\alpha\theta_1)760+(\mu_0-1)h_1},$$

hence if the pressure be adjusted to give a constant deviation, we have $m=m_1$, which determines $\nu_0$ as a function of $\mu_0$, a known quantity. If the exterior pressure and temperature have remained constant during the experiment, we have $\theta_1=\theta$ and $h_1=h$, and therefore

$$(\mu_0-1)h'=(\nu_0-1)h'_1,$$

or the refractive powers $(\mu-1)$ of two gases are inversely as the pressures required to produce the same deviation.

The refractive index of a gas may also be determined by delicate methods depending upon interference phenomena as mentioned in Arts. 122, 123.

INDICES OF REFRACTION OF GASES

| | Index. | Density. | | Index. | Density. |
|---|---|---|---|---|---|
| Air . . . | 1·000294 | 1·000 | Ethylene . . | 1·000678 | 0·978 |
| Oxygen . . . | 1·000273 | 1·106 | Marsh Gas . . | 1·000443 | 0·555 |
| Hydrogen . . | 1·000138 | 0·069 | Chloride of Ethyl . | 1·001095 | 2·234 |
| Nitrogen . . | 1·000300 | 0·971 | Hydrocyanic Acid. | 1·000451 | 0·944 |
| Chlorine . . | 1·000772 | 2·470 | Ammonia . . | 1·000385 | 0·596 |
| Nitrous Oxide . | 1·000503 | 1·520 | Phosgene . . | 1·001159 | 3·442 |
| Nitric Oxide . . | 1·000303 | 1·039 | Sulphydric Acid . | 1·000644 | 1·191 |
| Hydrochloric Acid | 1·000449 | 1·247 | Sulphurous Acid . | 1·000665 | 2·234 |
| Carbonic Oxide . | 1·000340 | 1·957 | Ether . . . | 1·001530 | 2·580 |
| Carbonic Anhydride | 1·000449 | 1·524 | Sulphide of Carbon | 1·001500 | 2·644 |
| Cyanogen . . | 1·000834 | 1·806 | Phosphuretted Hyd. | 1·000789 | 1·214 |

[Kayser and Runge using a Rowland grating and a prism of compressed air found the following indices corresponding to various wave lengths :—

INDICES FOR DRY AIR FOR 0° AND 760 MM.

| Wave Length. | Index. | Wave Length. | Index. |
|---|---|---|---|
| 5630 | 1·0002927 | 2860 | 1·0003088 |
| 4430 | 2955 | 2850 | 3094 |
| 4200 | 2967 | 2550 | 3158 |
| 3250 | 3035 | 2360 | 3219 |

These results they find agree very well with Cauchy's formula, and if $\lambda$ is expressed in millionths of a millimetre

$$\mu = 1{\cdot}00028817 + 1{\cdot}316\lambda^{-2} + 31600\lambda^{-4}$$

(*Astronomy and Astrophysics*, vol. xii. p. 426. See also Hasselberg, *ib.* p. 455).]

# CHAPTER VII

## INTERFERENCE FRINGES

**87. Destructive Interference.**—So far we have been engaged in considering the mode of propagation of a luminous wave, and the modifications which it undergoes when it encounters the surface of a new medium. We shall now proceed to inquire into the effects produced when two series of waves are propagated simultaneously in the ether from two small luminous origins close together.

When two waves arrive simultaneously at the same point of space, the ether there will be thrown into vibration by both, and we have already shown (chap. ii.) that in this compound motion each vibration may be regarded as acting independently of the other. If the constituent vibrations are in the same direction, the effects are added, and the amplitude of the resultant vibrations will be equal to the sum of the amplitudes of the constituents, but if they are opposed, the resultant amplitude is equal to the difference of the amplitudes of the constituents. In this latter case, if the vibrations are of equal amplitude, they should completely destroy each other. This is usually spoken of as destructive interference.

We know from experience that two sets of sound waves may neutralise each other and produce silence,[1] so also two sets of water waves when superposed may produce a dead level. If then there is any truth in the undulatory theory of light, something of the same kind should take place with two sets of light waves. That this actually occurs is abundantly proved by the following experiments.

Energy redistributed.

In all cases of interference, however, it is to be carefully remembered that light (regarded as energy) is never annihilated. The

[1] Two organ pipes tuned in unison and mounted close together produce only a very faint sound at external points. They start in opposite phases, and the effect which would be produced by one is just neutralised by the other. Other examples of interference occur in the beats of two notes nearly in unison, in the nodes of organ pipes, and vibrating strings, etc.

distribution alone is altered, so that the illumination, instead of being diffused regularly, is concentrated in some places at the expense of others.

**88. Two Small Apertures—Theory of Young's Experiment.**—Let us suppose that two sets of waves always exactly alike start from two near luminous origins A and B (Fig. 79). If the directions of the disturbances transmitted to any point P by the two sources conspire the amplitude of the disturbance at P will be doubled, but if the component vibrations be opposed at P they will destroy each other, and no effect will be produced at this point. In the former case the illumination at P is four times that produced by either of the sources acting singly; in the latter case the illumination is zero. The illumination sent by one source is swept away by that contributed by the other[1]—a result observed in the justly celebrated experiment of Dr. Young, and apparently opposed to the idea of the materiality of light (see p. 25).

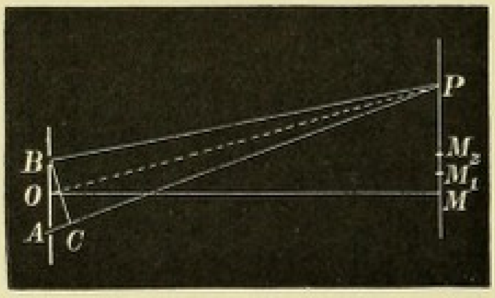



Fig. 79.

Let us now examine the theory of Young's experiment a little more closely. The direction of the vibrations sent by A and B to any point P of a screen will conspire, and the amplitude of the disturbance will be doubled, when they arrive at P in the same phase. Now we suppose the waves to set out from A and B in the same phase, so that they will arrive at P in the same phase if the path AP is equal to BP, or differs from it by any number of complete wave lengths. But the vibrations will be opposed at P if the waves arrive there in opposite phases, which will be the case if AP differs from BP by one half-wave length, or any odd number of half-wave lengths. We therefore conclude that if

$$AP - BP = n\frac{\lambda}{2},$$

the point P on the screen will be very bright, or dark, according as $n$ is even or odd. If then O be the middle point of AB, and if OM be perpendicular to AB, the distances AM and BM will be equal, therefore the lights should be in accordance at M, and it should be very bright. At $M_1$ there will be darkness if the difference of the distances of this point from A and B is half a wave. At $M_2$ there will be brightness again if the difference of its distances from A and B is two half-wave

[1] It is usually said that the light from one source is destroyed at these places by that from the other. This is misleading, however, as there is no real destruction, for the light is merely taken away from some places and heaped up at others.

lengths. Similarly, darkness will occur again at $M_3$, brightness at $M_4$, and so a series of bright and dark points occur alternately.

The point $M_n$ is such that the difference of its distances from A and B is $n\frac{\lambda}{2}$; if then this difference remains constant the point $M_n$ may lie anywhere on a hyperbola (as far as the plane of the paper is concerned), having A and B for foci.

In space the locus of $M_n$ is obviously an hyperboloid of revolution, viz. that generated by the revolution of the foregoing hyperbola round the line AB as axis.

On the screen then we have not a series of bright and dark points, but a series of alternately bright and dark lines, or bands perpendicular to the plane of the paper. These lines are the intersection of the screen with the hyperboloid loci just mentioned, which are so little curved as to sensibly coincide with their asymptotes.

The distance of any band from the central point M is very easily calculated. For if P corresponds to a retardation of $n$ half-wave lengths, the distance MP is small. Denote MP by $x$, MO by $a$, and with P as centre describe an arc of a circle BC. This arc is approximately a straight line perpendicular to OP, and AB is perpendicular to OM, therefore the angle ABC is equal to the angle POM. And hence their circular measures are equal, or

$$\frac{PM}{OM}=\frac{AC}{BC}, \text{ or } \frac{x}{a}=\frac{n\frac{1}{2}\lambda}{c},$$

where $c$ represents the distance AB. Hence

$$x=\frac{a}{c}\cdot n\frac{\lambda}{2}$$

and the point P is bright or dark according as $n$ is even or odd.

### *Examples*

1. The apparent angular distance from the centre M of any fringe of order $n$ as seen from the point O (Fig. 79) is independent of the distance of the screen MP.

[For if $\theta_n$ be the angle POM subtended by the central and $n$th bands at O we have

$$\theta_n=\frac{x_n}{a}=\frac{n\lambda}{2c}=\frac{\delta}{c}$$

which is independent of the distance $a$.]

2. The distance $x_n$ of the $n$th fringe from the centre M is proportional to the wave length and to the order ($n$) of the fringe and inversely as the apparent angle of the two sources as seen from M.

[For $x_n=\frac{a}{c}n\frac{\lambda}{2}$ and $\frac{c}{a}$ measures the apparent angle of the sources.]

**89. Colour and Wave Length.**—This formula shows that the distance of any fringe from the central one M depends on the wave length being in direct proportion to it. Hence, if composite light be used, we should expect to find rainbow-coloured bands[1] instead of merely bright and dark lines. This is what is actually observed, and, moreover, the inner edge of each band is violet, while its outer edge is red, showing that the violet wave lengths are shorter than the red.

Having accurately determined the magnitudes of $x$, $a$, $n$, and $c$, the formula gives the wave length of any particular kind of light. By this method the length of the red waves is found to be about ·0000266 in., the violet about ·0000167 in., and the mean wave about ·00002 or $\frac{1}{50000}$ in. or $\frac{1}{2000}$ mm.

**90. Fresnel's Mirrors.**—When Young first published his experiments, scientific men were by no means inclined to admit that the phenomena observed were due to interference in the manner conceived by their illustrious discoverer. It was known that the image of a small luminous origin, formed by the light admitted through a very small hole, was surrounded by coloured bands, and that light suffered a similar modification in passing near the edge of an opaque obstacle (see Newton's Observations, p. 227). The bands observed by Young might then be attributed to this modification (or diffraction). They might be a variety of diffraction bands.[2] Objections were therefore raised, and to remove them it was necessary to devise some method of obtaining two small sources of light close together wholly independent of apertures or edges of opaque obstacles. This was first contrived by Fresnel, whose experiments are justly ranked amongst the most important and instructive in the whole range of physical optics.

In his first experiment[3] Fresnel used two plane mirrors inclined at an angle of nearly 180°, so that they almost lie in the same plane. A beam of light diverging from the focus of a lens or from a very narrow slit is allowed to fall upon them. Each mirror reflects the light which falls upon it, and we have therefore two reflected beams whose directions are inclined at a very small angle.

If S (Fig. 80) be the source of light, OM and ON the two mirrors, the cone of light reflected from OM appears to come from a vertex A, the reflection of S in OM; similarly the light reflected from ON

[1] If the colour depends on the wave length. The existence of these bands consequently indicates that each coloured light has a definite wave length and period, just like each musical note.

[2] Diffraction bands are also due to interference and the principle of their production is essentially the same as that underlying interference bands. The latter are produced by the interfering action of two distinct waves, while the former are produced by interference between the elements of a single wave as explained in Art. 126.

[3] *Œuvres*, tom. i. pp. 150, 186, 268.

appears to come from B, the reflection of S in ON. If, therefore, the mirrors are inclined at a very obtuse angle, the points A and B will be very close together, and the reflected beams should give interference phenomena on a screen placed across any part of the region where they overlap, similar to those which would be produced if A and B were two small apertures. Here now we have two sources of light close together without the aid of edges or apertures, and the result is conclusive in favour of Young's theory. A brilliant system of fringes is produced, similar to those anticipated by the theory. In order to satisfy ourselves that these bands are really produced by the mutual action of the two beams, we have only to intercept one of them by covering the corresponding mirror with lampblack and the whole system instantly vanishes. They also vanish when the mirrors are parallel.

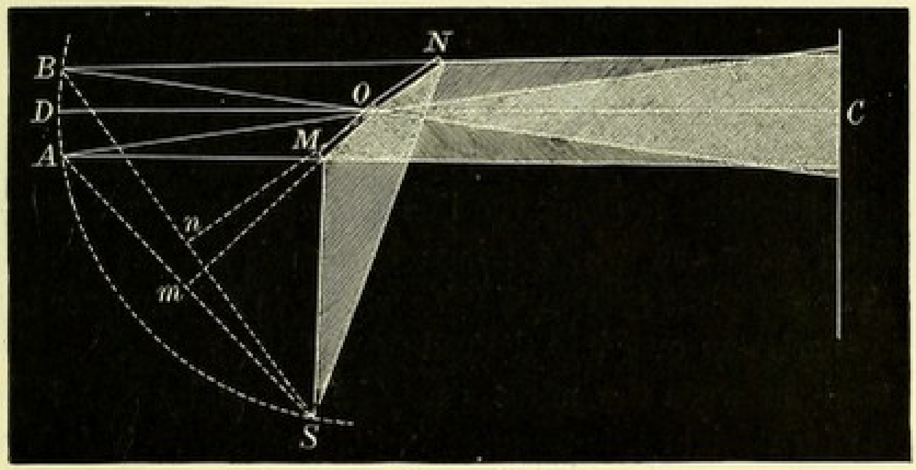



Fig. 80.—Fresnel's Mirrors.

Suppose the point of light S to lie in the plane of the paper, and let the line of intersection of the mirrors be perpendicular to it and meet it at O. Now SA is perpendicular to the mirror OM, and $Am = Sm$. Similarly BS is perpendicular to ON, and meets it produced at a point $n$ such that $Bn = Sn$, while the angle ASB is equal to the external angle between the mirrors, since it is the angle between the perpendiculars to them from S. Also since $OS = OA = OB$ the points S, A, B lie on a circle having its centre at O, therefore the angle ASB at its circumference is half the angle AOB at its centre, or the angle $AOB = 2\omega$ if $\omega$ denote the external inclination of the mirrors. Hence if DOC be perpendicular to both AB and the screen, D is the middle point of AB and C is the centre of the fringes. If we denote OD by $a$ and OC by $b$ we have $AB = 2a \sin \omega$, since OB is approximately equal to OD (or $\sin \omega = \tan \omega$), and the formula $x = \frac{CD}{AB} n \frac{\lambda}{2}$ for the distance of the $n$th fringe from the centre becomes

$$x=\frac{(a+b)}{2a\sin\omega}n\frac{\lambda}{2}=\frac{a+b}{2a\omega}n\frac{\lambda}{2},$$

since $\omega$ is a very small angle.

For the success of this experiment very careful adjustment is necessary. The polished surfaces of the mirrors should extend right up to the line of intersection of the two faces. If the mirrors are made of glass it should be black, or else silvered on the first face, otherwise the reflections from the second surface of the glass destroy the effect. Polished black glass is commonly used instead of polished metallic mirrors.

One mirror M is usually attached to a plate, which can be fixed to one of the uprights of the optical bench. The other mirror N can turn round the axis O; this axis is fixed to a plate through which three screws pass to adjust the level of the mirror M. Another screw is furnished with a spiral spring which keeps the mirror pressed against the three screws. By means of these the mirror M is adjusted till its edge is parallel to the axis O. When this is arrived at a screw enables the mirror N, by turning round O, to vary the angle between the mirrors. The other mirror can be screwed forward parallel to itself. This motion displaces A along the line SA, and the result is that the central band with the whole fringe system is displaced across the screen. The complete system of fringes may in this manner be caused to pass in succession over any desired part of the screen.

The angle between the mirrors may be found by first looking at the image of a straight line reflected in both mirrors. This image will be straight when the planes of the mirrors coincide, a position which is obtained by adjusting the mirror N till the image is straight. The number of turns of the screw which brings the mirror N from this position to any other measures the angle between the mirrors in the latter position. The angle between the planes of the mirrors may be also brought to zero by viewing the two images of the slit in the mirrors and adjusting N until the two images coincide. The number of turns of the screw required to effect this gives the angle between them. If the screw be not standardised the distance AB between the images may be found by the method of Ex. 2, p. 107.

The distances $a$ and $b$ can be measured on the scale of the optical bench; the distance $x$ by the micrometer motion of the cross wires in the eyepiece, and $n$ can be counted. We thus arrive at a determination of the wave length of any particular kind of light which we may choose to fix on.

In practice a narrow line of light (an illuminated slit) is used. The slit must be placed parallel to the line of intersection of the mirrors. In Fig. 80 the mirrors are planes perpendicular to the paper through OM and ON, while the slit is perpendicular to the paper through S.

**91. Fresnel's Bi-prism.** — In the above experiment a pencil of light was divided by reflection into two others inclined at a small angle and these produced the phenomena of interference. It is possible to procure the same result by refraction, and this is the basis of Fresnel's second experiment.[1]

Let CDE (Fig. 81) represent a glass prism with a very obtuse angle E, and let light from O fall perpendicularly on the opposite face CD. The whole prism is as if made up of two prisms CE and DE of very small angle (at C and D) placed base to base at E, and hence the name bi-prism.

The light which falls upon the upper half of the prism is bent downwards, and appears after emergence to diverge from a point B, while that which falls upon the lower half is bent upwards, and appears

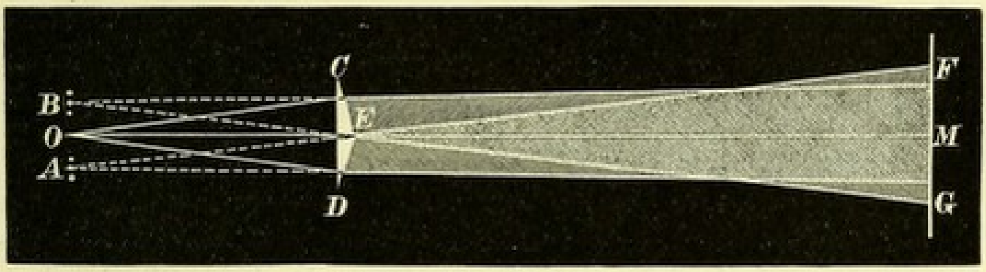



Fig. 81.—Fresnel's Bi-prism.

to diverge from A. The less the angles C and D the nearer together will be the points A and B, so that by diminishing these angles sufficiently the emergent cones of light will be similar to those coming from two very near origins, and interference effects will be presented as before.

The distances from the prism of the virtual foci A and B are very approximately the same as that of the luminous origin O. For since the refracting angles C and D are very small, the focal lines of each refracted cone coincide, and $\rho_1 = \rho_2 = \rho$ by Example 11, p. 115.

In practice a narrow strip of light, from a slit, is used. By this means very much increased brightness is obtained without loss of definition, as the various parts of the slit, if it be very narrow, give rise to coincident systems of bands. The length of the slit is carefully adjusted parallel to the edge of the prism, and the fringes are parallel to their common direction. The field is usually bordered with other

[1] This experiment has been sometimes wrongly attributed to Pouillet. It was first described by Fresnel (*Œuvres*, tom. i. p. 330).

systems of bands. These arise from diffraction, and will be explained farther on.

The distance of the $n$th band from the centre of the system is easily expressed. Thus if $a$ denote the distance of the origin O from the prism, and $b$ the distance of the prism from the screen, we have $AE = BE = a$ very approximately, and consequently if $c$ denote the distance AB between the virtual foci,[1] and $\delta$ the angle BEO or the deviation produced by the thin prisms, we have, if $\epsilon$ = angle of prism,

$$c = 2a \sin \delta = 2a(\mu - 1)\epsilon,$$

for since the angle of the prism is small $\sin \delta = \delta = (\mu - 1)\epsilon$. Hence

$$x = \frac{a+b}{c} n \frac{\lambda}{2} = \frac{a+b}{2a(\mu - 1)\epsilon} n \frac{\lambda}{2},$$

which shows that the bi-prism[2] is equivalent to a pair of mirrors inclined at an angle $(\mu - 1)\epsilon$.

**92. Peculiarities of the Bi-prism Fringes.**—The fringes produced by a bi-prism differ in some respects from those obtained by other methods. For, on account of the dispersion in the glass, the foci A and B will be different for the different colours. Thus the violet light will appear to come (on account of its greater refrangibility) from two points $A_v$ and $B_v$ a little farther apart than the two from which the

[1] An elegant method of determining $c$ is given by Prof. Glazebrook (*Phys. Optics*). Let $d$ be the distance from the prism to the focal plane of the eyepiece or screen where the fringes are depicted. Introduce a lens between the prism and eyepiece. This lens will form images of A and B in the focal plane if it is properly placed. Now in general two such positions of the lens can be found, and if $d_1$ and $d_2$, $d_1'$ and $d_2'$, be the distances of the lens from the focal plane and from the points AB in the two positions respectively we have

$$\frac{1}{d_1} + \frac{1}{d_2} = \frac{1}{f} = \frac{1}{d_1'} + \frac{1}{d_2'}.$$

But $d_1 + d_2 = d_1' + d_2'$, therefore $d_1 = d_2'$ and $d_2 = d_1'$. Hence if $c_1$ and $c_2$ be the corresponding distances between the images of A and B in the focal plane of the eyepiece we have, $c$ being the distance AB,

$$\frac{c_1}{d_1} = \frac{c}{d_2'}, \quad \text{and } \frac{c_2}{d_1'} = \frac{c}{d_2'}.$$

Therefore $$c_1 c_2 = c^2, \quad \text{or } c = \sqrt{c_1 c_2}.$$

$c_1$ and $c_2$ are measured by a micrometer eyepiece.

[2] If the prisms CE and DE were placed edge to edge instead of base to base, A and B would be interchanged and the emergent cones would have no common part. Interference bands could, however, be obtained by interposing a lens in the paths of the cones bringing them to real foci A′ and B′ images of A and B. The light diverging from A′ and B′ will produce fringes.

red appears to come. Denoting these distances by $c_v$ and $c_r$, the formula for the distance of the $n$th red band from the centre is (by Art. 88)

$$x_r = \frac{a+b}{c_r} n \frac{\lambda_r}{2},$$

and for the violet

$$x_v = \frac{a+b}{c_v} n \frac{\lambda_v}{2}.$$

The difference between $x_r$ and $x_v$ is consequently greater than if there was no dispersion by the prism. The iris-coloured bands are therefore broadened by the dispersion, and the overlapping is proportionately increased. The fringes of the bi-prism are bright, for the prism allows a great quantity of the light which falls upon it to pass through. These fringes are then bright, and very easily procured—the apparatus requiring very little trouble in setting up.

**93. Bi-plates.**—A beam of light may be subdivided by refraction through two plates, of the same nature and equal thickness, placed at an angle as indicated in Fig. 82. Two pieces M and N of parallel

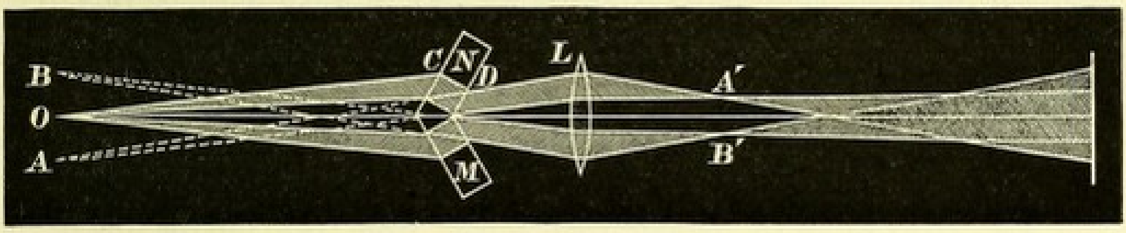



Fig. 82.—Bi-plates.

glass are cut from the same plate to ensure equality in thickness, and placed at an angle. On the bisector of the angle between them is placed the luminous origin O. The light which falls upon the plate N passes through it in a direction CD, and emerges parallel to its original direction, appearing to diverge approximately from a point A. Similarly the light which emerges from the plate M diverges from a virtual focus B. The emerging cones are received by a lens L which brings them to real foci A′ and B′. After diverging from A′ and B′ the beams will overlap and produce fringes.

The lateral displacement of a ray in passing through a parallel plate of thickness $e$ is easily seen to be $e \sin(i-r)/\cos r$, and consequently the distance $c$ between the virtual foci A and B is very approximately

$$c = 2e \sin(i-r)/\cos r,$$

which, if $2\theta$ denote the angle between the plates, may be put at once in the form

$$c = 2e \cos\theta[1 - \{\mu^2 + (\mu^2-1)\cot^2\theta\}^{-\frac{1}{2}}]$$

since $i = 90° - \theta$ very nearly.

**94. Lloyd's Single-mirror Fringes.**—A convenient method of displaying interference bands caused by the mutual action of direct and reflected light was devised by Dr. Lloyd[1] of Trinity College, Dublin. A polished mirror of metal or of black glass is placed so that the rays from a luminous origin B (Fig. 83), are reflected from it at nearly grazing incidence. The reflected rays diverge from a virtual focus A which is the image of the origin B, so that a point P on a screen placed beyond the mirror receives light directly from B, and also by reflection from the mirror, or, regarding A as the source of the reflected light, P is supplied by the origins A and B, which may be brought close together by making the angle of incidence nearly 90°.

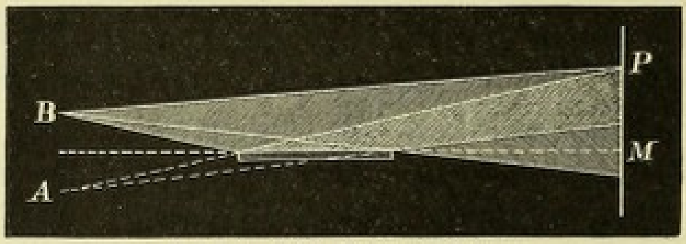


Fig. 83.—Lloyd's Single Mirror.

Since the reflected light is confined to the upper side of the mirror, less than half the complete system is formed, and it might be imagined that under no circumstances could more than one-half the system be obtained. However, by interposing a thin transparent plate in the path of the direct beam, or by holding the magnifier through which they are examined somewhat excentrically, the bands may be displaced (see Art. 102) so as to detach themselves from the mirror until the complete system is seen, as in Fresnel's experiments. The adjustments in this experiment are easily made; it requires no special apparatus, and the bands are bright and well marked.

Displaced centre.

Dr. Lloyd states that the centre of the system does not correspond to the line of intersection of the mirror and screen, but that the bands are all displaced through half the interval of a band width from the mirror edge. This, he suggests, indicates that the reflected light has been accelerated by half a wave length, or that its phase has been increased by $\pi$ at reflection.

**95. Fresnel's Three Mirror Experiment.**—Fringes produced by the use of three plane mirrors have also been obtained by Fresnel.[2]

[1] "A New Case of Interference of Rays of Light" (Lloyd, *Trans. Roy. Irish Academy*, vol. xvii.; read 27th January 1834).

Dr. Lloyd's method seems to have been anticipated by Professor Powell in the following passage: "Beautiful sets of colours (the theory of which is evidently dependent on interference) are seen on viewing a candle, or line of light, by very oblique reflection from any moderately polished surface, as ivory, ebony, etc., held close to the eye" (Rev. B. Powell, *Phil. Mag. and Ann.* January 1832). These bands may, however, have been due to diffraction.

[2] Fresnel, *Œuvres*, tom. i. p. 703.

Of the two pencils which produce the bands, one is reflected from the mirror M (Fig. 84) and the other successively from the mirrors L and N. The planes of the mirrors L and N intersect at O on the surface of M. Let $\omega$ and $\omega'$ be the angles they make with it respectively.

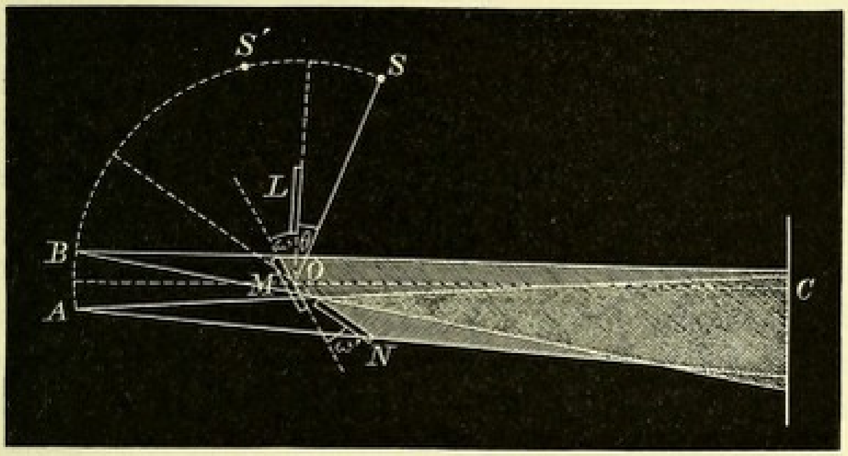



Fig. 84.—Fresnel's Three Mirrors.

Now if the line OS makes an angle $\theta$ with the mirror L, the light reflected from L will appear to come from a point S′ where the arc $SS' = 2\theta$, and if this light falls upon N its reflected beam will diverge from A, where the arc $S'A = 2(\omega + \omega' - \theta)$, for the mirror N makes an angle, $\omega + \omega' - \theta$ with OS′. Hence

$$SA = 2(\omega + \omega').$$

So also the light reflected from M will diverge from B where

$$SB = 2(\omega + \theta).$$

Consequently the arc

$$AB = 2(\omega' - \theta).$$

Hence if we denote OS by $a$ we have $OA = OB = a$, and the chord

$$AB = 2a \sin (\omega' - \theta).$$

The system is consequently equivalent to two mirrors inclined at an angle $\omega' - \theta$.

Since the mirror M is shaded by the others, the interfering pencils will be incomplete and unsymmetrical with respect to the central line OC. The central fringe may therefore be found near one extremity of the system, or it may lie entirely outside the visible fringes or common part of the interfering pencils. By increasing the path of the doubly reflected pencil, or diminishing that of the other, the fringes may be displaced on the screen so that the whole system may be viewed. This is readily done by screwing forward the central mirror M in the direction of the normal to its plane.

The angles $\omega$ and $\omega'$ may have any value up to 45°, for it is only necessary that $\omega' - \theta$ should be small. Fresnel worked with $\omega = \omega'$, and varied the angle between 7° 30′, and 40°.

If a change $\pi$ of phase accompanies each reflection the twice reflected beam will be accelerated by half a wave length on that reflected from the central mirror, and therefore, as in Lloyd's bands, the central band or that corresponding to equal distances CA and CB, will be black. This is confirmed by the experiment. It should also be further remarked that, as in case of a single mirror, the right side of A corresponds to the left of B and the left to the right, as if B were the reflection of A with respect to the central line OC (see Art. 98).

**96. The Optical Bench.**—Interference experiments are usually made on an optical bench. This apparatus usually consists of a horizontal bar, accurately graduated, along which three vertical uprights can slide freely. Attached to each upright is a vernier, so that the distance between them can be determined accurately by the scale on which they slide. Each is also furnished with a headpiece, which can be raised or lowered or turned round a vertical axis at will.

The first upright is furnished so as to hold a metal piece, which carries a slit capable of being adjusted to any convenient width by means of a screw. A fine adjusting motion of the head allows the slit to be brought accurately parallel to any desired direction. The second upright carries a frame in which can be placed a metal ring, which holds the plate containing the two small apertures, or any other apparatus for producing interference or diffraction, such as the bi-prism, a fine wire, an opaque edge, a diffraction grating, etc. The third carries a micrometer eyepiece furnished with a cross wire. The axis of this eyepiece is horizontal—that is, parallel to the bench on which all the pieces slide. The head of the second upright can be moved by means of a horizontal screw perpendicular to the length of the scale, so that each piece which it carries can be brought into the line joining the slit and observing telescope.

The use of a lens is legitimate in experiments on interference, for light brought to a focus by a lens is concentrated there without any relative change of phase in its components; since all the rays brought to that focus travel over paths which require the same time.

The eyepiece being furnished with a cross wire and micrometer screw, the distance $x$ of any band from the centre can be measured, and the value of $\lambda$ can be calculated accordingly.

**97. Conditions necessary for Interference.**—When commencing these investigations we assumed the waves emitted by A and B at any instant to be always exactly the same, and the theory indicates that this is necessary in order to have interference fringes or points at which the effects destroy each other *continually*. If the phases of the waves from A varied irregularly a great number of times per second with respect to those from B, we should have at any given point P neutralisation and co-operation succeeding each other so rapidly that nothing but a mean effect would be perceived, and this would be merely the sum of the mean effects of each source taken separately. Now if the disturbances come from two independent sources, such as the two different parts of a flame, the relative phases of the two would be purely casual, and no fixed and permanent neutralisation could be expected. Observation shows that no interference effects are manifested unless the two interfering streams of light come originally from the same source, and subsequently traverse slightly different paths, and this is what the theory anticipates.

In the experiments of Grimaldi the apertures were illuminated directly by the sun, and consequently no interference phenomena could possibly have been observed. This point was particularly noticed by Young, who allowed the sunlight to pass first through a narrow slit, and then through the two small apertures. He remarked that the fringes disappeared when one of the apertures was stopped, and also when the slit was removed, so that the two apertures were illuminated directly by the sun. In this case each point of the sun produces a distinct set of fringes, but the multitude of sets become so superposed and interspersed that all visible effect is obliterated.

An essential condition then is that the two apertures be supplied from the same source, so that the waves diverging from A and B at any instant may be exactly alike. To effect this in practice, a narrow slit is usually placed symmetrically near the apertures with its length perpendicular to the line joining them. The light from a lamp or other source falls first upon the slit, or is focussed on it by means of a condensing lens, and after diverging from it reaches the two apertures. The slit, being very narrow, is like a single line of light, each point of which is symmetrically situated with respect to A and B, and sends waves to each which are exactly alike, so that the whole resultant wave emitted by A is the same as that emitted by B. These waves, on arriving at any point P of the screen, will produce the phenomena of permanent interference.[1] The bi-prism and other apparatus for

[1] The fringes are only produced distinctly when the source is very narrow. The width of the aperture should be so small that the displacement of the centre of the system, incurred by using in turn the two edges of the slit as linear apertures, should be small compared with the width of an interference band.

producing interference bands are therefore to be regarded as contrivances to procure two similar origins of light in close proximity.

From this it will be easily understood how it is that two lamps, or two candles, can never be expected to destroy each other's effects anywhere when placed close together like two organ pipes tuned in unison. Two lamp flames have no permanent phase relation. The waves sent out by one are not necessarily similar to, or in any way related to, the waves sent out by the other. Each point of a flame is an independent source of light, and the waves emitted by it continually vary in character; while for two sources to produce darkness at any point the necessary condition is that they should *continually* send waves to this point opposite in phase, but in other respects exactly alike.

**98. The Corresponding Points of the Sources.**—In Young's experiment the two apertures are supplied by the same small source, and in the case of Fresnel's mirrors and bi-prism the interfering origins are images of the same source, and are therefore similar—the right-hand side of one corresponding to the right-hand side of the other and the left to the left. With Lloyd's single mirror it is somewhat different, for here the two interfering sources are the luminous origin and its image, but the right side of the image corresponds to the left side of the origin, and *vice versâ*. Now continuous interference can be expected to occur only between rays issuing from corresponding points of the interfering sources, for the waves emitted by the various points of the source (slit) have necessarily no fixed phase relation when it is supplied directly by a flame. Hence with the bi-prism and mirrors it is the right side of one image that interferes with the right of the other, and the left with the left, but in Lloyd's experiment the right side of the source interferes with the left of the image, and *vice versâ*. In the latter case when the slit has any sensible width the centre of symmetry must be the same for the bands produced by all corresponding points, but the distance $c$ between the corresponding points is variable. In Fresnel's experiments, on the other hand, the distance $c$ is constant, while the centre of symmetry varies. In Lloyd's experiment then the central bands are exactly superposed for all the groups of corresponding points, and the width of the slit does not interfere with the achromatism of the central line. The widths of the bands produced by different pairs of corresponding points will, however, be different (since the band width varies inversely as $c$), and this results in a confusion which increases from the centre outwards.[1]

[1] The measure of the confusion arising from the variation of $c$ in the finite width of the slit in Lloyd's experiment is easily found. Thus the distance of the $n$th

In Fresnel's experiments the band width is the same for all pairs of corresponding points, and the width of the slit merely leads to a lateral displacement of the central line of the various systems, so that the condition for distinctness is that the width of the slit be narrow compared with the width of a band, and this limiting width of the slit is independent of the order of the bands.

**99. Limit to the Number of Fringes.**—The formula (p. 141) for the distance of any bright band from the centre of the system shows that the width of any band, measured from darkness to darkness, is

$$w = a\lambda/c.$$

The band width is therefore directly proportional to the wave length of the light employed. If the light could be procured absolutely homogeneous—that is, of a single wave length $\lambda$—then theoretically the screen should be covered with an infinite number of similar bands, having nothing to distinguish one from another.

With ordinary light the case is very different. Each colour gives rise to a system of bands, and of these the red bands are broadest and the violet narrowest, the width of the former being about twice that of the latter. Hence it happens that after a few alternations the $n$th red band coincides with the $(n+1)$th violet, or perhaps the dark spaces of one system are filled up with the bright bands of another, so that overlapping and superposition of the multitude of systems from the red to the violet takes place, and this leads to the final obliteration of all visible effect at a short distance from the centre (Fig. 85). What occurs then is that a few (ten or twelve) bright rainbow-coloured bands are seen which become less and less distinctly marked, finally merging into one another and fading into uniform illumination at a short distance from the central line. This line is white, as it is a bright line for all wave

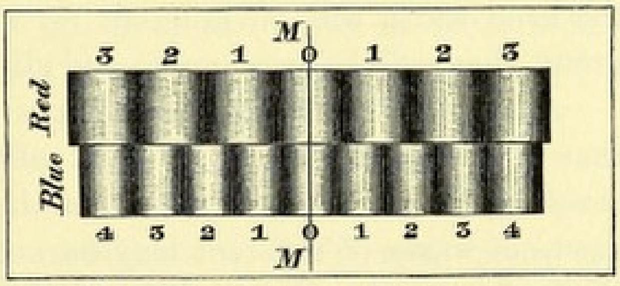



Fig. 85.—Overlapping of the Fringes.

band of any system from the centre is $x = na\lambda/c$, and therefore the interval $\delta x$ between two corresponding bands when $c$ varies is

$$\delta x = -\frac{na\lambda}{c^2}\delta c,$$

or if $w$ be the band width $= a\lambda/c$ we have

$$\frac{\delta x}{w} = -n\frac{\delta c}{c},$$

therefore the slit must be made narrower as $n$ increases if the distinctness of the bands is to be preserved.

lengths. Theoretically there is not a single place of complete darkness, for this would entail at that place a complete discordance of phase for all wave lengths, whereas any point at a distance $x$ from the centre will be bright, for the wave length determined by $na\lambda = cx$. The very existence of any visible bands with white light depends on the limited sensibility of the eye, which is confined to about one "octave," and on its capability of making chromatic distinctions.

Want of purity in the light is therefore detrimental to the production of a large number of visible bands.

**100. Interference under high Relative Retardation.**—It follows from the foregoing considerations that the more homogeneous the light the greater the number of observable interference bands, and if perfectly homogeneous light—that is, light of a single wave length—could be obtained, the screen should be covered with an infinite system of similar bands. In practice, however, it has not as yet been possible to obtain strictly homogeneous light. An approximation to it may be obtained by casting the spectrum of any source of light on a screen, furnished with a very narrow slit, in such a way that a narrow strip of the spectrum is transmitted through the slit. This slit may be used as the source of light in an interference experiment, and when so employed, the number of bands directly observable is vastly greater than that given by a source of ordinary white light. Still, in this case also, the light is heterogeneous, for the slit transmits a narrow band of the spectrum, and this band contains a *group* of waves varying in length by an amount depending on the width of the slit, and, as the band width will be different for the different constituents of the group, there will be overlapping, and ultimately obliteration of the bands.

A very convenient source of approximately homogeneous light is that of the sodium flame.[1] This consists of two narrow bands very close together in the yellow part of the spectrum, so that it contains a group of waves of different lengths, and overlapping ultimately occurs. With this source of light Fizeau observed 50,000 bands, and more recently this number has been largely increased by Professors A. A. Michelson and E. W. Morley.[2] Using the light of incandescent sodium vapour (in an exhausted tube provided with aluminium electrodes) interference was observed with a retardation of over 200,000 wave lengths, *i.e.* over 4 inches. This number was still further increased by using the light from Plücker tubes containing vapour of mercury or

[1] Proposed by Brewster, *Annales de Chimie et de Physique*, second series, tom. xxxvii. p. 437, 1828.

[2] A. A. Michelson and E. W. Morley, *Journal of the Association of Engineering Societies*, May 1888, and Address before the American Association, Cleveland Meeting, August 1888, by A. A. Michelson.

thallium chloride which gave interference with a difference of path of 540,000 and 340,000 wave lengths respectively. In these experiments a special form of apparatus, called an *interference refractometer*, was employed (Art. 123), by which any desirable difference of path could be easily introduced between the two interfering beams. Experiment consequently proves that the number of bands directly observable increases with the purity of the light.

It may also be shown that interference takes place in the regions beyond the limits of the visible fringes where overlapping exists to such an extent that the field appears to be uniformly illuminated. In these regions any point is a place of brightness for certain wave lengths and of darkness for others, and consequently the light there is a mixture in which only certain constituents of the original light are represented. Hence, if a slit be opened in any part of this region so that the light falling on it may be transmitted and examined in a spectroscope, the spectrum will exhibit certain dark bands corresponding to those waves which are destroyed by interference at the slit. This method was employed by Fizeau and Foucault.[1] If a narrow slit be opened[2] at the centre M of the fringe system the light which is transmitted through it will give a complete spectrum, since the central fringe is a place of brightness for all colours. On the other hand, if a slit be opened at any other part of the fringe system the transmitted light will give a spectrum exhibiting those colours for which the slit is a place of brightness, and consequently crossed by dark bands showing the absence of those colours which are destroyed by interference at the slit. If $x$ be the distance of the slit from the centre of the system, the wave lengths of the dark bands satisfy the equation

Spectroscopic analysis.

$$x=\frac{a}{c}(2n+1)\frac{\lambda}{2}.$$

In the experiment of Fizeau and Foucault a slit was opened at the centre of the system produced by Fresnel's mirrors, and by screwing forward one of the mirrors parallel to itself, the system of fringes was displaced gradually across the screen, so that they passed in succession over the slit. As the first dark band comes on the slit, a dark band is seen in the spectrum which passes from the violet to the red as the mirror is gradually displaced, then a second succeeds it, and then two, three, or more dark bands simultaneously appear crossing the

[1] Fizeau and Foucault, *Annales de Chimie et de Physique*, third series, tom. xxvi. p. 138, 1849; *Comptes Rendus*, 24th November 1845.

[2] The width of the slit should not be more than a moderate fraction of that of a band.

spectrum, till finally they become so numerous and narrow that to separate and distinguish them the resolving power of the spectroscope requires to be increased. Hence the extent to which interference can be observed is limited only by the resolving power of the spectroscope.[1]

If two dark bands appear in determinate parts of the spectrum corresponding to wave lengths $\lambda$ and $\lambda'$, then by the above equation

$$(2n+1)\lambda=(2n'+1)\lambda'=2\delta,$$

where $n$ and $n'$ are two whole numbers, and $\delta$ is the difference of the distances of the interfering origins from the slit in the screen. Between these two bands a number N of other dark bands may occur which can be counted. Then we have

$$\mathrm{N}+1=n'-n=\frac{2n+1}{2}\frac{\lambda-\lambda'}{\lambda'},$$

or

$$2n+1=(\mathrm{N}+1)\frac{2\lambda'}{\lambda-\lambda'},\quad 2n'+1=(\mathrm{N}+1)\frac{2\lambda}{\lambda-\lambda'}.$$

Knowing $n$ or $n'$ we can calculate $\delta$, the relative retardation of the pencils when they reach the slit.

**101. Achromatic Interference Bands.**—The interference bands

[1] It has been usually held that even if the number of observable fringes were not limited by overlapping, yet a major limit to this number would be determined by certain irregular changes taking place in the source of light. Thus if we consider a point P so far from the central line of the fringes, that AP − BP is a large number of wave lengths, then the light reaching P at any instant from B is a large number of wave periods in advance of that which reaches it from A. The former wave was emitted from the source some time before the latter, and if the nature of the waves emitted by the source has changed in the meantime, the waves reaching P at any instant from A and B will be dissimilar, and have no constant phase relation.

The *irregularity* contemplated in this view of the subject does not appear to have any clearly defined meaning in the case of white light. For example, in the case of a monochromatic source, if the vibration be represented by a simple equation of the form $x=a\sin(\omega t+\alpha)$, we deal with an infinite train of waves propagated in a perfectly regular manner. If this train be supposed broken up by sudden changes of phase originating in the source, then we have to deal with a system of groups of waves, and the conditions of propagation will be altered, so that the state of affairs at any point ceases to be represented by the foregoing simple equation, and the problem becomes much more complicated. In M. Gouy's opinion the nature of white light may be best understood by assimilating it to a disturbance originated by a sequence of entirely irregular impulses. The action of a prism is to analyse this complex disturbance into its constituents in the way a complex periodic function is analysed into its simple harmonic components in a Fourier series. That interference bands may be observed under high relative retardation with white light is merely a proof of the resolving power of the spectroscope and affords no criterion as to regularity in the vibration (see further, M. Gouy, *Journal de Physique*, 2$^{me}$, tom. v. p. 354, 1886; Lord Rayleigh, *Phil. Mag.* vol. xxvii. p. 460, 1889; Arthur Schuster, *Phil. Mag.* vol. xxxvii. p. 509, 1894).

ordinarily obtained are highly coloured, and this happens because, in the formula,

$$x=\frac{a}{c}n\frac{\lambda}{2},$$

the distance $c$ is the same for all colours while $\lambda$ is variable.[1] If, however, by some means $c$ be made different for the different colours so as to be directly proportional to $\lambda$, we will have $\lambda/c=$ const., and the bands will be of the same width for all colours. The fringes will therefore be achromatic, and the want of homogeneity of the light will offer no obstacle to the production of a large number of fringes. This may be easily arranged by using Lloyd's mirror and a diffraction grating (Art. 135, etc.) with which to form a spectrum. White light from a narrow slit falls in succession upon a grating and an achromatic lens, so as to form diffraction spectra in the focal plane of the lens. One of these spectra[2] is used as the proximate source of light in the interference experiment, and since the deviation of any colour in the diffraction spectrum varies as $\lambda$, it is only necessary to arrange the mirror so that its plane passes through the white central image in order to realise the conditions for achromatic bands. When the adjustments are carefully made the whole field is filled with fine bands, which become coloured only at the edges of the field.

With less perfection the diffraction spectrum may be replaced by a prismatic one so arranged that $\lambda/c$ is constant for the most luminous rays. "The bands are then achromatic in the same sense that an ordinary telescope is so. In this case there is no objection to a merely virtual spectrum, and the experiment may be very simply executed by Lloyd's mirror and a prism of about 20° held just in front of it."

"It is interesting to observe the effect of coloured glasses upon the distinctness of the bands. If the achromatism be in the green, a red or orange glass, so far from acting as an aid to distinctness, obliterates all the bands after the first few. On the other hand, a green glass, absorbing rays for which the bands are already confused, confers additional sharpness. With the aid of a red glass a large number of bands are seen distinctly, if the adjustment be made for this part of the spectrum."[3]

**102. Displacement of the Fringes—Diminished Speed in Denser Media.**—In the deduction of the law of refraction the wave theory

[1] In the case of the bi-prism $c$ varies, being least for the red and greatest for the violet, and this exaggerates the overlapping and increases the colouring.

[2] Lord Rayleigh used the second.

[3] Lord Rayleigh, *Phil. Mag.* vol. xxviii. pp. 77, 189, 1889.

pointed out that the velocity of light should be less in the denser or more refracting media, and we mentioned in passing that the emission theory pointed to an opposite conclusion. This point can now be decided by means of the phenomena of interference. It is obvious that the central fringe is situated in that place to which it takes the light the same time to travel from the two interfering origins. If now a thin plate of glass (or other transparent substance) be interposed in the path of one of the beams, the light of that beam will be retarded or accelerated according as it travels slower or faster in the glass than in air. The point then at which the two beams will arrive in the same time will be displaced on the screen. The central band will be moved towards the path of the beam in which the plate (Fig. 86) is interposed if the light travels slower in the glass, but to the opposite side if it travels quicker in it. The result is decisively in favour of the diminished velocity of light in the more refracting media, and the difference of velocity may be measured by the amount by which the fringe is displaced.

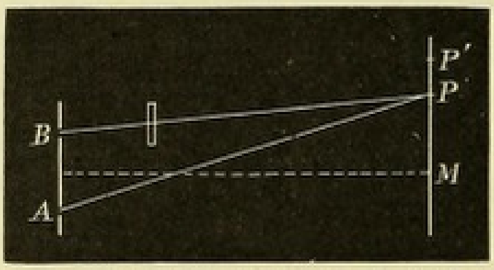


Fig. 86.—Displacement of the Central Band.

It is easy to calculate the relation between the displacement and the refractive index of the interposed plate. Let P be the position of the central fringe when the plate is interposed in the path of the ray BP. The time of travelling over AP is the same as that of travelling over BP; hence if $v$ and $v'$ denote the velocities in air and in the plate respectively, and $e$ the thickness of the plate, then

$$\frac{\mathrm{BP}-e}{v}+\frac{e}{v'}=\frac{\mathrm{AP}}{v},$$

or

$$\mathrm{BP}-e+\mu e=\mathrm{AP}.$$

Hence

$$e(\mu-1)=\mathrm{AP}-\mathrm{BP}=n\frac{\lambda}{2},$$

if the central fringe is displaced through the distance occupied by $n$ fringes, a result which was obvious, for the retardation introduced by the passage through the plate is $(\mu-1)e$.

**103. Application to the Determination of Refractive Indices.**—The amount of displacement in the foregoing important experiment furnishes us with a method of determining the refractive index of a substance when the displacement of the central fringe and the thickness of the plate are known.[1] Fresnel and Arago[2] applied it to the determination of the refractive indices of gases. It is susceptible of great accuracy, the minutest change in the index of refraction of air being observed,—such, for instance, as the change due to a rise of $\frac{1}{20}$ of a degree in temperature.

By the same method it was ascertained that the refractive index of dry air is about one-millionth greater than that of air saturated with vapour. Arago also pointed out that the scintillation of the stars is due to interference arising from the changes in the refracting powers of portions of the atmosphere through which the different portions of light reach the eye.

M. Billet[3] has devised a very convenient method of producing interference fringes, and showing the effect of a plate interposed in the path of one of the interfering pencils in displacing the fringes. It consists of a lens cut into two halves, L and L′ (Fig. 87), which can be separated or brought close together at will by means of a micrometer

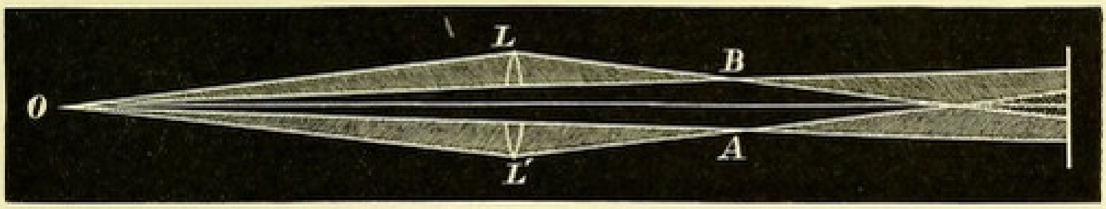



Fig. 87.—Billet's Split Lens.

screw. Their sections are brought parallel by turning one of them round a fixed axis by means of another screw. The luminous origin O produces two images A and B situated on the optic axes OA and OB of the two halves, the motion of which allows the images A and B to be brought as close together as desired. The light diverging from these images produces fringes on a screen placed at any part of their common path, and it is easy to interpose plates of any transparent substance in the path of either or of both simultaneously.

With the bi-prism the thin plate may be placed over one-half of the prism, or if two plates are to be compared, one may be placed over each half. These experiments were first executed by Fresnel and Arago,[4] and gave rise to the construction of interference refractometers.

[1] It also determines $e$ when we know $\mu$ and the displacement.

[2] By observing the position of the fringes formed by two rays, one of which passed through vacuum and the other through air.

[3] *Ann. de Chimie et de Phys.*, third series, tom. lxiv. p. 385.

[4] *Œuvres de Fresnel*, tom. i. pp. 125, 691, and several memoirs of Arago (*Œuvres*, tom. x. pp. 298, 312).

An ingenious modification of Billet's experiment has been recently suggested by M. G. Meslin,[1] in which fringes of a circular form are obtained. Thus in all the forms of experiment so far described the line AB joining the interfering sources is at right angles to the direction of propagation of the light, and for this reason the fringes obtained on the screen on which the light falls are approximately right lines parallel to AB. This is the case because the surfaces of constant retardation are hyperboloids of revolution round the line AB, having A and B for foci. But if by any means the line AB is turned so as to be parallel to the direction of propagation of the light—that is, if it is turned so as to pass through O—then the axis of revolution of the surfaces of interference will be perpendicular to the screen, and their cross sections on the screen will be a system of concentric circles. In Billet's experiment AB is at right angles to the central line, because L and L′ are separated by displacement at right angles to this line, but if the displacement were made parallel to this line, then the line AB would pass through O.

Circular fringes.

This is shown in Fig. 88, in which the lower half of the lens is displaced through an interval CC′ parallel to the central line. With this arrangement the light which passes through L is brought to a focus A on the line joining O to its optic centre C, while the light passing through L′ is brought to a focus B on the same line, and if a screen PQ be placed anywhere between A and B, interference fringes will be depicted on it in the region of the overlapping beams of light. The surfaces of constant retardation in this case are not hyperboloids but ellipsoids of revolution round the line AB, and having A and B for foci. For if we consider any point X on the screen which receives light from both L and L′, the light reaching it from L will have travelled over a path $\delta + AX$, where $\delta$ represents the equivalent path from O to A and is the same for all rays passing through L. Similarly the light reaching X from L′ will have travelled over a path $\delta' - BX$, where $\delta'$ is the equivalent path from O to B through L′. The difference of path at X for the two constituents is consequently

$$(\delta + AX) - (\delta' - BX) = (\delta - \delta') + AX + BX.$$

But
$$\delta' - \delta = AB = \text{const.},$$

consequently the path retardation at X is

$$AX + BX - AB,$$

[1] G. Meslin, *Comptes Rendus*, 1893; and *Journal de Physique*, 3$^{me}$, tom. ii. p. 205, 1893.

and if this is to remain constant we must have

$$AX + BX = \text{const.}$$

That is, X must lie on the surface of an ellipsoid of revolution, having A and B for foci. The cross section of any one of these ellipsoids by the plane PQ is a circle, and bright and dark bands are accordingly a

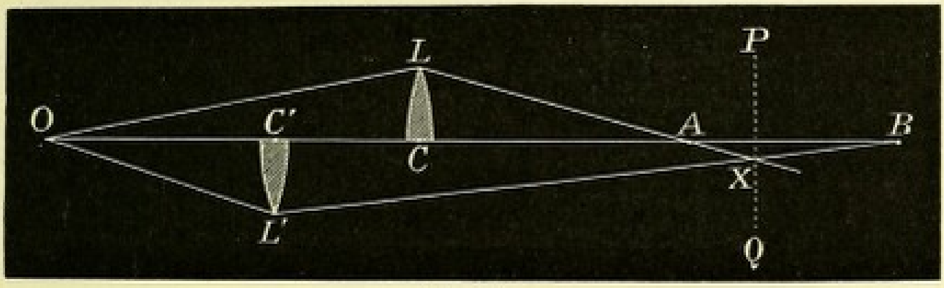



Fig. 88.

system of concentric circular rings. These circles are not, however, complete, for the beams of light overlap only on one side of the central line, and at most only half of the complete system, *i.e.* semicircles can be obtained.

### *Example*

For a given position of the screen prove that the radii of the consecutive rings are proportional to the square roots of the whole numbers 1, 2, 3, etc.

[Take the middle point of AB as origin and the line AB as axis of $x$, so that the equation of any one of the ellipses may be written in the form

$$\frac{x^2}{\alpha^2} + \frac{y^2}{\beta^2} = 1,$$

where $AX + BX = 2\alpha$. Hence if $AB = 2c$ the relative retardation at X is

$$\delta = AX + BX - AB = 2(\alpha - c).$$

But $\beta^2 = \alpha^2 - c^2$, and consequently, since $\delta$ is small compared with either $\alpha$ or $c$, we have approximately

$$\beta^2 = (\alpha - c)(\alpha + c) = \tfrac{1}{2}\delta(\alpha + c) = c\delta\,;$$

since $\alpha = c + \frac{1}{2}\delta$ is nearly equal to $c$. Hence, in the equation of the ellipse,

$$y = \frac{\beta}{\alpha}\sqrt{\alpha^2 - x^2},$$

we may write $\alpha = c$ and $\beta = \sqrt{c\delta}$, so that we have

$$y = \sqrt{\frac{\delta}{c}(c^2 - x^2)}.$$

But $y$ is the radius of the ring corresponding to the retardation $\delta$, and for the bright and dark rings $\delta = \frac{1}{2}n\lambda$, according as $n$ is even or odd.

When the screen passes through the middle point of AB the rings reach their maximum size, when $x = c$ they degenerate to zero.]

**104. Abnormal Displacement of the Central Band.**—It has been pointed out by Sir G. G. Stokes[1] that the method of determining the refractive index of a plate by the displacement of a system of interference fringes is subject to a theoretical error depending on the dispersive power of the plate. In the absence of dispersion the retardation ($\delta$) introduced by the plate would be independent of $\lambda$, and would therefore be completely compensated at the point $x = a\delta/c$. But when there is dispersion the retardation $\delta$ depends on $\lambda$, and the different colours are unequally retarded by the plate. The violet fringe system will consequently be most displaced, and the red least. If $u$ be the linear displacement of the fringe system of wave length $\lambda$, we have $u = a\delta/c$ and $\delta = (\mu - 1)e = f(\lambda)$ suppose, consequently

$$u = \frac{a}{c}f(\lambda) \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (1)$$

The centre of the complete displaced system is therefore not necessarily at the point reached by the two pencils in the same time, but is determined by the coincidence of bright bands of the most brilliant parts of the spectrum. Measured from the original centre, the position of the $n$th bright band of wave length $\lambda$ will be

$$x = \frac{a}{c}n\lambda + u = \frac{a}{c}\left\{n\lambda + f(\lambda)\right\} \quad . \quad . \quad . \quad . \quad . \quad (2)$$

When this quantity is as independent of $\lambda$ as possible, the best coincidence of the various bright bands will occur, and the position will correspond to the centre, or, as Cornu terms it, the achromatic band of the displaced system. This will happen when $dx/d\lambda = 0$, or when $n$ is the nearest integer to

$$n = -f'(\lambda) = -\frac{c}{a}\frac{du}{d\lambda}.$$

Substituting this value of $n$ in (2) we find for the displacement of the central white band

$$x = \frac{a}{c}\left\{f(\lambda) - \lambda f'(\lambda)\right\} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (3)$$

where $f(\lambda) = (\mu - 1)e$.

This when expressed in terms of $u$ gives by (1)

$$x = u - \lambda\frac{du}{d\lambda} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (4)$$

[1] *Brit. Assoc. Report*, 1850; also *Math. and Phys. Papers*, vol. ii. p. 361.

and when expressed in terms of the band width $w = a\lambda/c$, we have Airy's[1] formula

$$x = u - w\frac{du}{dw} \quad . \quad . \quad . \quad . \quad . \quad . \qquad (5)$$

The final term on the right-hand side of each of these equations is inherently negative, for since the refrangibility increases as the wave length decreases, it follows that the displacement of the bands corresponding to a given wave length must increase as the wave length (or band width) decreases. Hence $du$ and $d\lambda$ (or $dw$) must have opposite signs. The final term in the foregoing equations (which represents the abnormal displacement of the central white band caused by dispersion) will consequently be additive, and will *increase* the normal effect of the interposed plate.

*Exercise.*—Assuming the truth of Cauchy's formula,

$$\mu = A + \frac{B}{\lambda^2} + \frac{C}{\lambda^4} + \ . \ . \ .$$

the relative error is

$$\frac{\lambda f'(\lambda)}{f(\lambda)} = -\frac{1}{\mu - 1}\left(\frac{2B}{\lambda^2} + \frac{4C}{\lambda^4} + \ . \ . \ .\right),$$

$$= -\frac{2}{\mu - 1}\left(\mu - A + \frac{C}{\lambda^4}\right).$$

**105. Abnormal Displacement of the Fringes by a Prism — Potter's Experiment.** — Repeating an experiment of Professor Powell's,[2] Mr. Potter[3] found that a prism interposed in the pencils of light reflected from Fresnel's mirrors, arranged to produce interference bands, systematically deviated the fringes towards the thick end of the prism by an amount greater than that of the calculated centre of the interfering pencils. This he considered inconsistent

[1] "Remarks on Mr. Potter's Experiment on Interference" (G. B. Airy, *Phil. Mag.* March 1833).

[2] Powell's experiment was a simple variation of that of Art. 102. The interposition of the thin plate is attended by difficulties on account of the extreme proximity of the interfering pencils. Powell suggested the use of a thin prism of 4° or 5° refracting angle, the edge of the prism being parallel to the line of intersection of the mirrors. The two pencils then pass through different thicknesses of the prism. Powell says, "The whole set of stripes are seen in the deviated image, unaltered, except by a trifling degree of colour and a slight shifting towards the more refrangible end of the spectrum, obviously due to prismatic refraction" (Rev. Baden Powell, *Phil. Mag. and Ann.* January 1832).

[3] The prism used by Potter was an ordinary glass prism of 43° angle (R. Potter, jun., *Phil. Mag.* February 1833).

with both the wave and emission theories, but Sir G. B. Airy[1] showed at once that this phenomenon followed as an immediate consequence of the wave theory. The investigation is that which we have already applied to determine the position of the central or achromatic band when the fringes are displaced by a plate interposed in the path of one of the pencils. If $u$ be the linear displacement of the fringes corresponding to a wave length $\lambda$ and $w$ the width of a band, the displacement of the achromatic band is

$$u - w\frac{du}{dw}$$

as before. The quantity $du/dw$ being negative, the abnormal effect is added to the regular deviation produced by the prism.

The abnormal displacement of the central band is therefore a consequence of the heterogeneity of the light. In fact, with perfectly homogeneous light we would have nothing whereby to distinguish the central band, for the fringe system would consist of a set of perfectly similar bands. The effect of the prism is to displace the apparent centre of the system. The $n$th band is rendered achromatic, but the system is no more achromatic than before, for the widths of the component bands and the overlapping remain unaltered.

If a diffraction grating be used instead of a prism, the deviation will vary as the wave length—that is, $u$ varies as $w$, and consequently $u - wdu/dw = 0$.

**106. Talbot's Bands.**—A remarkable system of bands was discovered by H. F. Talbot,[2] and their complete explanation was first given by Airy,[3] whose calculation is very complicated, but his final result may be obtained from very elementary considerations, which are given in the theory of diffraction (see Art. 155). At present we shall merely give Talbot's general account. He describes his experiment as follows: "Make a circular hole in a piece of card of the size of the pupil of the eye. Cover one-half of this opening with an extremely thin film of glass (probably mica would answer the purpose as well, or better). Then view through this aperture a perfect spectrum formed by a prism of moderate dispersive power, and the spectrum will appear covered throughout its entire length with parallel obscure bands resembling the absorption produced by iodine vapour."

In Fig. 89 the thin plate is represented in three different positions.

[1] "Remarks on Mr. Potter's Experiment on Interference" (G. B. Airy, *Phil. Mag.* March 1833).

[2] H. F. Talbot, *Phil. Mag.* 1837, part i. p. 364.

[3] Airy, *Phil. Trans.* 1840, part ii. p. 225, and 1841, p. 1.

At L it is situated directly between the eye of the observer and the eyepiece of the telescope, at M it is placed between the prism and the object glass, and at N between the prism and the collimator.

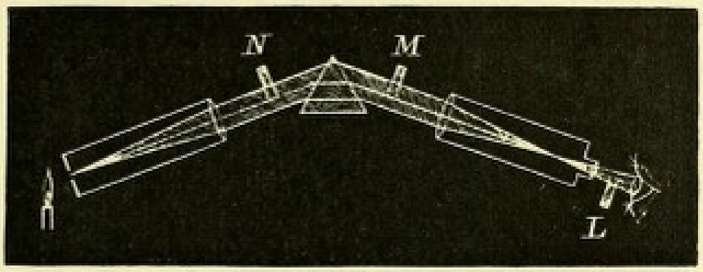


Fig. 89.—Talbot's Bands.

An imperfect explanation of the bands was given by Talbot on the principle of simple interference. Thus if $\delta$ be the retardation suffered by any ray, of wave length $\lambda$, in passing through the plate, then one-half of the light passing through the aperture will be retarded relatively to that which passes through the other, and if the quotient $\delta/\lambda$ is a whole number, the two halves will agree in phase, but if $\delta$ is equal to $(2n+1)\frac{1}{2}\lambda$, they will be opposite in phase and destroy each other. Now $\delta/\lambda$ varies from one colour to another, so that agreement and opposition in phase will recur alternately as we pass from one end of the spectrum to the other. It is consequently traversed by a system of dark bands.

A similar experiment is that of Brewster,[1] by which he imagined he had discovered what he termed a new polarity of light. "While examining the solar spectrum formed in the focus of an achromatic telescope after the manner of Fraunhofer, he placed a thin plate of glass before his eye in such a manner as to intercept and retard one-half of the pencil which was entering one-half of the pupil. He was then surprised to find that when the edge of the retarding glass plate was turned towards the red end of the spectrum, intensely black lines made their appearance . . . but upon turning the plate of glass half round (still keeping its plane perpendicular to the axis of the eye), so as to present the edge past which the rays entered the eye to the violet end of the spectrum, the dark bands disappeared." In intermediate positions the bands appeared more or less distinct, according as the edge was more presented to the red or to the violet end. The thinner the glass the more distinct the lines, and they were formed in any part of the spectrum. Brewster remarked that "An examination of these lines affords the very best means of determining the dispersive powers of substances; for their distance from one another increases or diminishes exactly as the entire length of the spectrum is increased or diminished, and the number of them in the same part of two spectra of different lengths is always the same."

[1] "On a new Polarity of Light" (Sir D. Brewster, *Brit. Assoc. Report*, 1837).

**107. Powell's Bands.**[1]—In a hollow glass prism (Fig. 90) containing some highly refractive liquid, such as oil of sassafras, anise, or cassia, a plate of glass is inserted with its lower edge parallel to the edge of the prism, and so that its plane nearly bisects the angle of the prism, while it extends only through the upper half of the liquid, leaving the lower or thinner part clear. The light from a slit being transmitted through it in the usual manner, the spectrum thus formed is crossed by a number of dark bands parallel to the slit or edge of the prism.

With some liquids and plates the bands are sensibly equidistant; in others they increase in number and fineness towards one end of the spectrum. In most cases they extend throughout, but in some they are deficient at one part of the spectrum.

If the thickness of the plate exceed a certain limit the bands become too numerous and too fine to be seen; if less than a certain

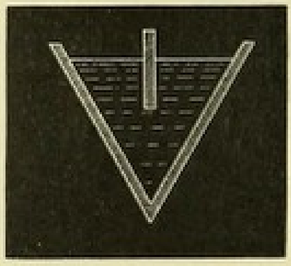


Fig. 90.—Powell's Bands.

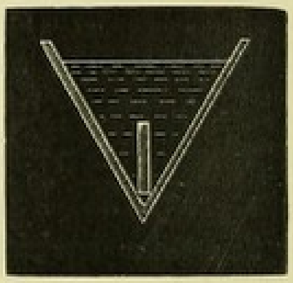


Fig. 91.

limit they become too few, broad and faint; while for some intermediate thickness they appear most vivid and distinct.

When the plate is inclined either way, even to being in contact with the side of the prism, the bands are still seen, but they suffer a slight displacement downwards as the plate is inclined.

Some combinations of liquid and plate, such as glass with oil of turpentine, or water, give no bands with this arrangement. Stokes pointed out that in this case bands are produced by placing the plate in the thinner part of the prism, leaving the wider part clear (Fig. 91).

With plates of crystallised substances, such as Iceland spar, two sets of bands, one finer than the other, are presented. On applying a Nicol's prism (chap xi.) each set disappears alternately, leaving the other visible at each quarter of a revolution of the analyser, showing that they are due to the two oppositely polarised pencils. The finer bands belong to the extraordinary, and the broader to the ordinary ray.

[1] "On a new Case of Interference of Light" (Baden Powell, *Phil. Trans.* 1848).

The explanation of the general[1] formation of the bands is afforded by the simple interference theory. The plate having a refractive index differing from that of the liquid, causes one part of the pencil passing through the prism to be retarded relatively to the other, by an amount which increases from one end of the spectrum to the other; and as this difference of phase amounts successively for the various colours to an odd or even multiple of $\pi$, the two parts will be in discordance for some or accordance for other waves, and produce corresponding dark or bright bands in the spectrum.

These phenomena are analogous to those observed by Fox Talbot and Sir D. Brewster, on partially intercepting the spectrum by a plate of mica covering half the pupil of the eye. Here the retardation is that due to the difference of refraction of a plate of glass and an equal thickness of liquid, and in the other cases it is the difference between the mica and the displaced air. Hence Stokes has varied the experiment by inserting the glass plate in a vessel with parallel sides, and allowing the light from a prism to fall upon it (Fig 92).

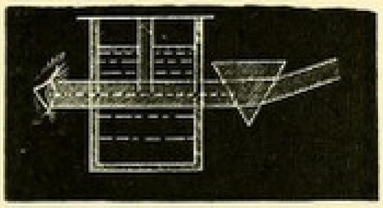


Fig. 92.—Stokes's Modification.

## *Examples*

1. Assuming[2] the truth of Cauchy's law of dispersion, determine to what degree of approximation $\lambda/c$ can be made independent of $\lambda$ by means of a prism in the experiment of Art. 101.

[According to Cauchy's law we may take $c = A - B/\lambda^2$, and therefore

$$\frac{\lambda}{c} = \frac{\lambda^3}{A\lambda^2 - B}.$$

Hence if $\lambda/c$ is to be stationary when $\lambda$ has a prescribed value $\lambda_0$, we must have $d(\lambda/c) = 0$, that is,

$$A\lambda_0^2 = 3B.$$

The proportional deviation of $\lambda/c$ from the prescribed value $\lambda_0/c_0$ is consequently

$$\frac{\lambda/c}{\lambda_0/c_0} = \frac{2}{3\lambda_0} \times \frac{\lambda^3}{\lambda^2 - \frac{1}{3}\lambda_0^2}.]$$

2. If the achromatism is to hold good in the neighbourhood of the D line ($\lambda_D = \cdot 58890$), find the proportional variation of $\lambda/c$ for the C line ($\lambda_C = \cdot 65618$).

[Using the formula of Ex. 1 we have $\lambda_0 = \cdot 58890$ and $\lambda = \cdot 65618$, therefore

$$\frac{\lambda/c}{\lambda_0/c_0} = 1 \cdot 0155.]$$

---

[1] The complete investigation has been given by Stokes (*Phil. Trans.* p. 227, 1848; *Math. and Phys. Papers*, vol. ii. p. 14).

[2] These Examples are taken from Lord Rayleigh's article, *loc. cit.*

3. In Ex. 2 determine the order ($n$) of the band at which the C-system is displaced through half a band width relatively to the D-system.

[The relative displacement at the $n$th bright band is $\delta x = na(\lambda/c - \lambda_0/c_0) = \frac{1}{2}a\lambda_0/c_0$, if the displacement is half a band width.

Hence

$$n = \frac{\frac{1}{2}\lambda_0/c_0}{\lambda/c - \lambda_0/c_0}.$$

Using Ex. 2 this reduces to $n = 32$—that is, after 32 complete periods the bright bands of one system coincide with the dark bands of the other.]

4. If the prism be not employed, prove that the bright bands of one of the systems of Ex. 3 will coincide with the dark bands of the other when $n = 4{\cdot}2$.

[When the prism is not employed $c$ is constant and the formula for $n$ in Ex. 3 becomes

$$n = \frac{\frac{1}{2}\lambda_0}{\lambda - \lambda_0} = 4{\cdot}2.]$$

5. If the two systems differ in wave length by a small amount $\delta\lambda$, prove that the formula for $n$ in Ex. 3 becomes approximately

$$n = \frac{1}{3}\left(\frac{\lambda_0}{\delta\lambda}\right)^2\left[1 + \frac{7}{3}\frac{\delta\lambda}{\lambda_0}\right].$$

[This formula gives the order of the band at which complete discrepance first occurs between the systems $\lambda_0$ and $\lambda_0 + \delta\lambda$, and it shows that when $\delta\lambda$ is small the order of the band is inversely proportional to the square of $\delta\lambda$.

The corresponding effect will occur without a prism at the band

$$n = \frac{\frac{1}{2}\lambda_0}{\lambda - \lambda_0} = \frac{\frac{1}{2}\lambda_0}{\delta\lambda},$$

so that the effect of the prism is to increase the number of bands in the ratio

$$2\lambda_0 : 3\delta\lambda.]$$

# CHAPTER VIII

## INTERFERENCE BY ISOTROPIC PLATES

### Section I.—The Colours of Thin Plates

**108. General Statement of the Phenomena.**—The examples of interference which we are now about to discuss are noteworthy on account of the peculiarities which they present and their frequent occurrence to ordinary observation. Here it is no longer necessary to have a very narrow source of light.

When ordinary white light falls upon a thin film of a transparent substance, such as a soap bubble or a film of oil spread on the surface of water, brilliant colours are generally observed, and sometimes all the tints of the rainbow are exhibited. These colours were first observed by Boyle and Hooke, and the latter succeeded in blowing glass sufficiently thin to exhibit them distinctly.[1] They are often developed in mica and other minerals which possess a lamellar structure, but the most familiar instance of their exhibition is in the froth of liquids and films of oil. The colours vary with the thickness of the film, and disappear altogether when it exceeds certain limits. This is well exhibited by dipping the mouth of a wine glass into soap water. The viscid aqueous film which adheres to it after immersion displays the whole succession of these phenomena. When the film covering the mouth of the glass is held in a vertical plane it appears at first uniformly white, but as it grows thinner by the gradual descent of the fluid, colours begin to be exhibited at the top, where it is thinnest. These colours form horizontal bands which become more and more brilliant as the thickness diminishes, but when the thickness is reduced to a certain limit at the upper part the film becomes quite black, and it has at this place arrived at such a stage of tenuity that it is no longer able to support its own weight, and the film bursts.

Every one is familiar with the fact that polished steel becomes

[1] A novice in the art of glass-blowing may succeed in this experiment.

coloured in various shades when exposed to the air. These colours are due to a thin film of the oxide of the metal which is gradually formed on the surface.[1]

The same appearances are displayed in a still more striking manner when two plates of glass (which contain a thin film of air between them) are pressed together, or when a convex lens is laid on a plate of glass. Around the point of nearest approach successions of coloured rings of great brilliancy are presented, which dilate as the pressure is increased so as to diminish the thickness of the included air film.

**109. Thin Plates—Retardation.**—Let us now consider the case of homogeneous light falling upon a thin uniform plate or film of a transparent isotropic substance, for example, a film of air enclosed between two parallel plates of glass.

The light incident in the direction $A_1B_1$ (Fig. 93) on the first surface of the film is divided there into two portions, one reflected parallel to BC and the other transmitted parallel to $B_1C_1$. This latter portion is further divided at the second surface, one part being transmitted and the other reflected back along $C_1B$ to suffer refraction at B and emerge in part from the plate again in the direction BC, making an angle with the normal equal to the angle of incidence.

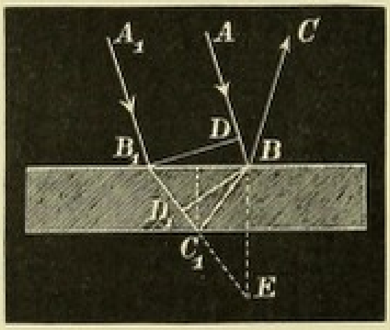


Fig. 93.—Thin Plate, Retardation.

The direction BC is therefore the same as that of the light which is directly reflected at the first surface. In this direction we have therefore two streams of light, one coming from the first surface by the reflection there of the incident light, and the other emerging from the first surface after it has been refracted into the plate and reflected at the second surface. This latter stream of light having traversed the plate, will be retarded relatively to the light which is reflected directly at the first surface. The two streams will reinforce or weaken each other according as they are in the same or in opposite phases.[2]

[1] This oxide is formed rapidly when the temperature is high, and the thickness of the film depends so invariably on the temperature that artists are in the habit of estimating the temperature by the colour developed. Thus steel in the process of tempering is spoken of as having received a yellow heat or blue heat, etc.

[2] "This mode of explaining the phenomena of thin plates was pointed out by Hooke in a remarkable passage in his *Micrographia* some years before the subject was taken up by Newton. In this passage he very clearly describes the manner in which the rings of successive orders depend on the interval of retardation of the second 'pulse' or wave with respect to the first, and therefore on the thickness of

The calculation of the difference of phase of these two streams of light is very simple. Draw $B_1D$ and $BD_1$ perpendicular to AB and $B_1C_1$ respectively. Thus $B_1D$ and $BD_1$ are the fronts of the incident and refracted waves respectively. At the instant the light from D reaches B the disturbance from $B_1$ has travelled to $D_1$ in the second medium. Accordingly the first pencil of light is about to leave the surface at B along BC at the same instant as the refracted ray is leaving $D_1$ to traverse $D_1C_1B$. The retardation is therefore

$$\delta = C_1D_1 + C_1B.$$

Produce $B_1C_1$ to meet at E the normal drawn to the plate at B. Then if $r$ be the angle of refraction into the plate it is obvious that $BEC_1 = r$, $C_1B = C_1E$, and therefore

$$\delta = D_1E = BE \cos r = 2e \cos r,$$

where $e$ is the thickness of the film.

So far the theory seems to indicate that the brightness will be greatest when the difference of path $2e \cos r$ is an even number of half wave lengths, and least when this difference is an odd number of half waves, but the results of observation show that the conditions of light and shade are exactly reversed.

Now if the difference of phase depends only on the difference of path traversed by the pencils, when this difference vanishes (which is the case when the thickness of the film is infinitesimally small) the two pencils should be in the same phase and the illumination should be a maximum. But we know that if at any point the thickness of the plate is zero there will not be any light reflected there, and the point will appear dark; the light passes straight through. Hence we have arrived at two opposite conclusions, and of these two the latter is undoubtedly correct. The difference of phase therefore must depend on something else as well as on the difference of path traversed by the pencils. This second element is not far to seek.

It has been shown (Art. 57) that when a pulse travels along a row

the plate. But he does not seem to have had any distinct idea of the principle of interference itself, and his conception of the mode in which the colours resulted from this 'duplicated pulse' is entirely erroneous. Euler was the next who attempted to connect the phenomena of thin plates with the wave theory of light, but the attempt, like all the physical speculations of this great mathematician, was signally unsuccessful, and the subject remained in this unsettled state until the principle of interference was discovered by Young. When this principle was combined with the suggestion of Hooke the whole mystery vanished. The application was made by Young himself, and all the principal laws of the phenomena were readily and simply explained" (Lloyd, *Elementary Treatise on the Wave Theory of Light*, third edition, p. 138).

of balls (or a cord) and passes into another row of different density, it is subdivided into two pulses, one traversing the second system and the other reflected in the first system. Let us confine our attention to the reflected wave. Suppose the first pulse to be propagated as a forward displacement, then the reflected pulse will be propagated as a displacement in the direction of propagation or the reverse, according as the first system is less or more dense than the second. In the first case the pulse is reflected without change of sign, in the second it is reflected with change of sign. The displacement in one case is in the opposite direction to that in the other, so that if the two were superposed they would neutralise each other. This is expressed by saying that the waves are in opposite phases (see further, Art. 112).

Change of sign.

Now in the case of the thin plate, if the light at the upper face is reflected in passing from a dense medium to a rare, then at the second surface it is reflected in passing from rare to dense. The two portions are reflected under opposite conditions, so that if one is reflected without change of sign, the other is reflected with change of sign, and the two reflected waves are in opposite phases. Accordingly the act of reflection under the opposite conditions introduces half a period difference of phase. Hence the whole retardation is

$$2e \cos r + \tfrac{1}{2}\lambda$$

measured in the medium of which the plate is composed. If $\mu$ be the refractive index of the material of the plate, the air distance which corresponds to this is

$$2\mu e \cos r + \tfrac{1}{2}\lambda$$

where $\lambda$ is now the wave length in air. The illumination of the plate will consequently be a maximum if $2e \cos r =$ an odd number of half waves (measured in the plate), a minimum if $2e \cos r =$ an even number of half wave lengths. If $e = 0$, or the thickness of the plate is infinitely small, the retardation is half a wave and the illumination is a minimum, which agrees with experiment. If the wave length be measured in air the brightness will be greatest when $2\mu e \cos r$ is an odd number of half wave lengths, and least when the same quantity is an even number.

In the foregoing we have been considering homogeneous light—that is, light of a definite wave length $\lambda$. If, however, the incident light is heterogeneous, and contains many wave lengths for which $\lambda$, $\mu$, and $r$ are different, it will happen that the condition for brightness will be satisfied by some of the constituent waves, while the condition for darkness is satisfied by others. It follows, therefore, that in the

White light.

reflected beam only some of the constituents of the original light will be represented.

Thus when ordinary solar light is incident on a thin film the light which comes from any point of it to the eye will not include any of the wave length satisfying the equation ($\lambda$ being the wave length in air)

$$2\mu e \cos r = n\lambda$$

where $n$ is any whole number. The light from this point will be accordingly coloured, the colour at any point depending both on the thickness of the film and on the angle of incidence. If the angle of incidence or the thickness varies from point to point of the film, corresponding variations of colour will occur, but if the incidence and thickness be constant, the colour will be uniform. If the light returning from any point of the film be analysed in a spectroscope the spectrum will consequently be crossed by certain dark bands corresponding to those waves which have been neutralised by interference. The number and closeness of these dark bands will increase with the thickness of the plate, and when the thickness reaches a certain limit they become so fine and close that the resolving power of the spectroscope may fail to separate them. This means that interference has ceased to be observable by reason of excessive overlapping. This overlapping increases with the thickness of the plate, and leads to the obliteration of the bands, but it should be remembered that interference takes place in thick plates just as in thin, in the same way as interference exists in Fresnel's experiment in the regions distant from the central line. The thinner the plate the less the overlapping, and the more observable the phenomenon. This overlapping will be diminished, and the limiting thickness of the plate at which interference can be observed will be increased by increasing the homogeneity of the light.

Overlapping.

A simple case is that in which the film is a thin wedge bounded by two planes inclined at a small angle. The lines of equal thickness are parallel to the edge of the film, consequently with monochromatic light it will appear crossed by a system of parallel bright and dark bars, and with white light these are replaced by a system of parallel coloured bands.

In general, however, it is not necessary to resort to any particular contrivance in order to obtain interference fringes by reflection from thin films. If a film of uniform thickness could be procured it would appear uniformly coloured when a beam of parallel light is reflected from it. In practice, however, it is impossible to secure perfectly plane surfaces, and the film of air enclosed between two so-called

Fringes.

parallel plates of glass is not of uniform thickness throughout. As a consequence the film, instead of being uniformly coloured, generally exhibits coloured fringes forming curves which encircle the points of nearest approach of the plates. We are thus furnished with an exceedingly delicate optical test of the planeness of the surfaces. On pressing the plates closer together so as to reduce the thickness of the film the bands dilate, showing how the colour at any point depends upon the thickness of the film.

**110. Influence of Dispersion on the Colour of a Film — Condition for Achromatism.**—The order of the colours in fringes presented by thin films (or the series of tints passed through by a film of uniform thickness as the thickness is varied) may differ considerably from that exhibited in the interference bands produced by the mirrors of Fresnel and Lloyd. This arises from the fact that in the latter case the path retardation ($\delta$) at any given point is the same for all wave lengths, and consequently the phase retardation $\delta/\lambda$ varies from colour to colour by reason of the variation of $\lambda$ alone. In the case of a film, however, there is a second agency through which the phase retardation may be altered, namely, the dispersion within the film, and the change arising from this cause may be either of the same or of opposite sign to that arising from the variation of $\lambda$ when there is no dispersion, so that the colour effect may be either increased or diminished by it.

Thus, let us consider a definite case in which a uniform film, enclosed between two infinite media A and B, is viewed by an eye situated in A, and let a parallel beam of white light traversing the medium A fall upon the film so as to be reflected to the eye in question. In this case the angle of incidence is the same for all wave lengths, but the angle of refraction into the film is different for the different colours. As a consequence the path retardation $2e \cos r$ varies from colour to colour, and the expression for the phase retardation for a given colour (namely, $2e \cos r/\lambda$) varies both in its numerator and denominator. When the variation of $\cos r$ is the same as that of $\lambda$ for all values of $\lambda$, the phase retardation will be the same for all colours—that is, all the colours will be equally affected by interference, and the dispersion in the film will have produced achromatism. This will happen when

$$\frac{\cos r}{\lambda} = \text{constant},$$

and the corresponding angle of incidence is determined by this condition, where $\lambda$ is the wave length measured in the film.

It can be easily seen that achromatism may be produced in this manner when the film is less refracting than the media between which it is enclosed. For in this case the angle of refraction into the film increases with the refrangibility of the light—that is, $r$ increases, or $\cos r$ diminishes, as $\lambda$ diminishes, and consequently there may be some angle of incidence for which the variation of $\cos r$ is the same as that of $\lambda$ from colour to colour, and $\cos r/\lambda$ may be the same for all. When the film is more highly refracting than the enclosing media, on the other hand, the value of $\cos r$ diminishes as $\lambda$ increases, and that the effect of dispersion in the film is to exaggerate the colour effect which would be produced by the variation of $\lambda$ alone.

At nearly perpendicular incidence the variation of $\cos r$ for the different colours is very small, and the effect of dispersion on the colour of the plate is not very sensible, but as the angle of incidence approaches the angle of total reflection from the film, the angle of refraction approaches $90°$, and the variation of $\cos r$ with $\lambda$ becomes much more considerable, and may even pass the limit at which colour compensation is produced. If perfect compensation should occur, the bands produced by a film of variable thickness will be simply black and white, and a uniform film will pass through alternations of black and white as its thickness is varied.

In order that perfect achromatism may be produced it is clear that the refractive index of the film must be related to the wave length in some particular manner—that is, there must be a particular law of dispersion in the film. This law is contained in the achromatic condition $\cos r/\lambda = \text{const.}$ Thus, if we write $\cos r = k\lambda$, we have

$$\sin^2 r = 1 - k^2\lambda^2,$$

and consequently if $i$ be the angle of incidence (which in the foregoing case is the same for all colours) at which achromatism occurs, we have $\sin r = \mu \sin i$, where $\mu$ is the refractive index of the film with respect to the medium A, and consequently the law of dispersion required is

$$\mu^2 \sin^2 i = 1 - k^2\lambda^2,$$

which may be written under the more general form

$$\mu^2 = a + b\lambda^2.$$

It should be observed that the common case of a film of air enclosed between two parallel plates of glass does not satisfy the conditions demanded by the foregoing investigation unless some special contrivance be adopted to cause the light to enter the glass so as to fall in a parallel beam on the film.[1] If no such precautions be taken a parallel beam of white light falling upon the first glass plate will be dispersed within the plate, so that the angle of incidence on the film will vary from colour to colour, while the angle of refraction into the film will be the same for all colours, and equal to the incidence on the glass plate. In this case the light falling upon the film is not a parallel beam of white light, but a dispersed beam, and the dispersion which occurs in the glass is corrected by an equal and opposite dispersion in the film, so that the light within the film is a parallel beam of white light. The path retardation $2e \cos r$ is consequently the same for all colours, and the phase retardation $2e \cos r/\lambda$ varies to the full extent from colour to colour without compensation of any sort. The series of tints passed through by such a film, as its thickness is varied, should therefore be the same as that presented in the interference fringes produced by Fresnel's mirrors.

### *Example*

If the refractive index $\mu$ be related to the wave length $\lambda$ by the equation

$$\mu = a + \frac{b}{\lambda^2} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \qquad (1)$$

prove that achromatism will be most nearly approached when[2]

$$\cot^2 r = \frac{2(\mu - a)}{\mu}. \quad . \quad . \quad . \quad . \quad . \quad . \qquad (2)$$

[Colour compensation will take place when $\cos r/\lambda$ is stationary—that is, when

$$\lambda \sin r dr + \cos r d\lambda = 0 \quad . \quad . \quad . \quad . \quad . \quad . \qquad (3)$$

But $\mu \sin i = \sin r$, therefore since $i$ is the same for all colours, we have $\sin r/\mu =$ constant, and consequently

$$\mu \cos r dr - \sin r d\mu = 0 \quad . \quad . \quad . \quad . \quad . \quad . \qquad (4)$$

Combining (3) and (4) we obtain

$$\cot^2 r = -\frac{\lambda}{\mu}\frac{d\mu}{d\lambda} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \qquad (5)$$

Now by equation (1) we find

$$\frac{d\mu}{d\lambda} = -\frac{2b}{\lambda^3} = -\frac{2}{\lambda}(\mu - a),$$

consequently (5) reduces at once to the required relation (2).

---

[1] This may be easily arranged by using a prism and a plate (as in Fig. 99) instead of two plates. A parallel beam of white light incident normally on one of the faces of the prism will fall on the film as a parallel beam of white light, and it may be so arranged that the reflected light falls normally on the second face of the prism so as to emerge to an eye in air in the condition in which it leaves the film.

[2] Lord Rayleigh, *Phil. Mag.* vol. xxviii. p. 192, 1889; and *Ency. Brit.* Art. "Wave Theory."

In the case of Chance's "extra-dense flint" glass

$b = \cdot 984 \times 10^{-10}$, and for the sodium lines $\mu = 1 \cdot 65$,
$\lambda = 5 \cdot 89 \times 10^{-5}$, consequently achromatism occurs when
$r = 79^\circ \cdot 30'$.]

**111. More Complete Investigation—Multiple Reflections.**—The theory of thin plates as it came from the hands of Young laboured under an imperfection which, however, was soon removed. Thus it is easily seen that the intensities of the two portions of light reflected from the two surfaces of the plate are not equal.[1] These two portions therefore can never wholly destroy one another, and the intensity of the light in the dark rings can never entirely vanish, as it appears to do when homogeneous light is employed. Poisson was the first to point out and to remedy this defect in the theory. It is evident, in fact, that there must be an infinite number of partial reflections within the plate, at each of which a portion is transmitted, and it is the sum of all these portions that must be taken into account.

In the foregoing discussion we assumed that the only light which emerged from the plate along BC (Fig. 94) is that ray which after refraction at $B_1$ is reflected at $C_1$. It is obvious, however, that there is a multitude of other rays which also emerge in this direction. For if we take $BB_1 = B_1B_2 = B_2B_3$, etc., it is clear that a ray incident at $B_2$ will after refraction pass along $B_2C_2B_1C_1B$ and part of it will emerge at B along BC. Similarly a ray incident at $B_3$ will, after alternate reflections at the two surfaces of the plate, emerge in part along BC. It is thus evident that the complete stream of light which issues along BC from the interior of the plate consists of several parts, the first of which is by far the most powerful and the others diminish rapidly to zero. The foregoing calculation is therefore only approximate, and it becomes necessary to calculate each of the components and to sum their effects. If the retardation suffered by the ray $A_1B_1C_1$ in the plate is $\delta$, we have $\delta = 2e \cos r$, and it is obvious that the retardations of the consecutive rays incident at $B_2$, $B_3$, $B_4$, etc., are $2\delta$, $3\delta$, $4\delta$, etc., respectively. We thus know the phases of the components as they arrive at B, but to calculate their joint effect it is necessary also to

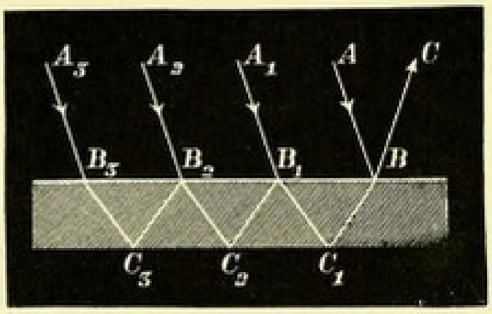


Fig. 94.—Multiple Reflections.

[1] The condition for the equality of these two portions would violate the condition of Art. 112.

know their amplitudes. For this purpose it is necessary to examine the relation connecting the reflection and refraction coefficients.

**112. Relation connecting the Reflection and Refraction Coefficients.**—If the amplitude of the incident ray AB be $a$, then the amplitude of the reflected ray BC may be denoted by $ab$ where $b$ is a proper fraction; similarly the amplitude of the refracted ray may be denoted by $ac$ where $c$ is some other proper fraction.[1] Thus if a ray of amplitude $a$ be incident at B along AB (Fig. 95), it will give rise to two rays BC and BD of amplitudes $ab$ and $ac$ respectively, and if these rays be reversed they should combine again to give a ray along BA of amplitude $a$. But if we reverse BC it will give us two rays, one reflected along BA of amplitude $ab^2$, and another refracted along BE of amplitude $abc$. Now in reversing the ray BD we cannot suppose that the amplitudes of its reflected and refracted components are obtained by multiplying its own amplitude by $b$ and $c$ respectively, for it is passing from the lower to the upper medium, whereas the ray AB passes from the upper to the lower. We will therefore suppose that the components of BD reversed, along BA and BE, are $acf$ and $ace$ respectively. We must then have

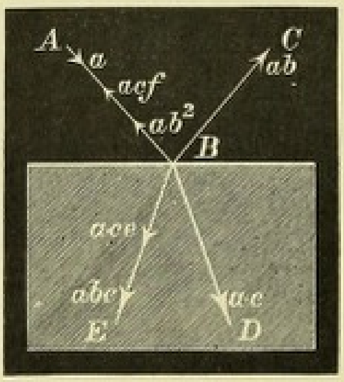


Fig. 95.—Reflection and Refraction Coefficients.

$$acf + ab^2 = a,$$

and

$$ace + abc = 0,$$

since the two rays BC and BD reversed must give a ray along BA alone of amplitude $a$. The above equations give

$$cf = 1 - b^2 = 1 - e^2,$$

and

$$b = -e.$$

The last equation, if the incidence is $i$ in one medium and $r$ in the other, shows that the amplitude of the ray arising from reflection at the surface in passing from the upper medium to the lower is equal to the amplitude of the reflected ray which would arise if the same ray were reflected in passing from the lower medium to the upper, but of opposite sign. The reflection under the different conditions therefore changes the phase of the vibration by half a period. There is half a

Change of sign.

[1] This investigation was given by Stokes ("On the perfect Blackness of the central Spot in Newton's Rings," *Cambridge and Dublin Math. Jour.* 1849, vol. iv. p. 1; or *Mathematical and Physical Papers*, vol. ii. p. 89).

wave lost by one relatively to the other. Beyond the critical angle we have $f=0$, and $e=1$.

If the amplitude of the light incident at B is $a$, and its phase $\phi$, the amplitude of the reflected vibration is $ab$, and its equation is

$$y=ab\sin\phi,$$

if we assume that no change of phase is introduced by the act of reflection,[1] while the equation of the vibration reflected in passing in the reverse direction DB—that is, from the second medium to the first—is

$$y=-ab\sin\phi.$$

*Cor.*—Using the energy equation of Art. 68 we obtain at once

$$\frac{\rho}{\rho'}\frac{1-b^2}{c^2}=\frac{\sin 2r}{\sin 2i}.$$

Hence by the equation $(1-b^2)=cf$, we have

$$\frac{f}{c}=\frac{\rho'}{\rho}\frac{\sin 2r}{\sin 2i}$$

as the relation connecting the refraction coefficients $c$ and $f$.

Using Fresnel's form of the energy equation, we have

$$\frac{f}{c}=\frac{\tan i}{\tan r}=\frac{\mu\cos r}{\cos i},$$

and using MacCullagh's, we have

$$\frac{f}{c}=\frac{\sin 2r}{\sin 2i}=\frac{\cos r}{\mu\cos i}.$$

**113. Calculation of the Intensity.**—We can now calculate the intensity of the light reflected from the plate.

The amplitude of the light which emerges at B (Fig. 94) after incidence at $B_1$ is $acef$, and its phase is $\phi+\delta$; similarly the amplitude[2] of the light which emerges from B along the path $B_2C_2B_1C_1B$ is $ace^3f$, and its phase is $\phi+2\delta$. The amplitude of the next ray is $ace^5f$ and its phase is $\phi+3\delta$. The equations of these vibrations are $y_1=acef\sin(\phi+\delta)$, $y_2=ace^3f\sin(\phi+2\delta)$, $y_n=ace^{2n-1}f\sin(\phi+n\delta)$. Their sum is, writing $-b$ for $e$,

$$y=-\frac{acf}{b}[b^2\sin(\phi+\delta)+b^4\sin(\phi+2\delta)+b^6\sin(\phi+3\delta)+\ .\ .\ .\ ].$$

[1] Stokes shows that the same relations hold between $a$, $c$, $e$, $f$, if reflection and refraction are accompanied by a change of phase, p. 183 *infra*.

[2] Sir G. Airy, *Trans. Camb. Phil. Soc.* vol. iv. p. 419, 1830.

The expression inside the bracket may be written by expansion in the form

$$P \sin \phi + Q \cos \phi,$$

where

$$P = b^2 \cos \delta + b^4 \cos 2\delta + b^6 \cos 3\delta + \ . \ . \ .$$

and

$$Q = b^2 \sin \delta + b^4 \sin 2\delta + b^6 \sin 3\delta + \ . \ . \ .$$

Therefore writing $\cos n\delta$ in the form $\frac{1}{2}(e^{\iota n\delta} + e^{-\iota n\delta})$, we find

$$P = \tfrac{1}{2}(b^2 e^{\iota\delta} + b^4 e^{2\iota\delta} + b^6 e^{3\iota\delta} + \ . \ . \ . \ ) + \tfrac{1}{2}(b^2 e^{-\iota\delta} + b^4 e^{-2\iota\delta} + \ . \ . \ . \ )$$

$$= \tfrac{1}{2}\left(\frac{b^2 e^{\iota\delta}}{1 - b^2 e^{\iota\delta}}\right) + \tfrac{1}{2}\left(\frac{b^2 e^{-\iota\delta}}{1 - b^2 e^{-\iota\delta}}\right) = b^2 \frac{\cos \delta - b^2}{1 - 2b^2 \cos \delta + b^4}.$$

Similarly

$$Q = \frac{b^2 \sin \delta}{1 - 2b^2 \cos \delta + b^4}.$$

Now the equation of the resultant displacement due to the two streams of light along the direction BC is

$$y = ab \sin \phi - (P \sin \phi + Q \cos \phi)\frac{acf}{b}$$

$$= X \sin \phi + Y \cos \phi$$

where

$$X = ab - abcf \frac{\cos \delta - b^2}{1 - 2b^2 \cos \delta + b^4} = \frac{2ab(1 + b^2) \sin^2 \frac{1}{2}\delta}{1 - 2b^2 \cos \delta + b^4}$$

since $cf = 1 - b^2$, and similarly

$$Y = \frac{-abcf \sin \delta}{1 - 2b^2 \cos \delta + b^4} = \frac{-ab(1 - b^2) \sin \delta}{1 - 2b^2 \cos \delta + b^4}.$$

Hence

$$X^2 + Y^2 = \frac{4a^2b^2 \sin^2 \frac{1}{2}\delta}{(1 - 2b^2 \cos \delta + b^4)^2}\{(1 + b^2)^2 \sin^2 \tfrac{1}{2}\delta + (1 - b^2)^2 \cos^2 \tfrac{1}{2}\delta\}$$

$$= \frac{4a^2b^2 \sin^2 \frac{1}{2}\delta}{1 - 2b^2 \cos \delta + b^4}.$$

Now the vibrations $y = X \sin \phi$, and $y = Y \cos \phi$ denote two whose difference of phase is 90°, therefore the amplitude of the resultant vibration is $\sqrt{X^2 + Y^2}$, or the intensity of the resultant light in this case is measured by $X^2 + Y^2$. Hence the expression for the intensity is

$$I = \frac{4a^2b^2 \sin^2 \frac{1}{2}\delta}{1 - 2b^2 \cos \delta + b^4} = \frac{4a^2b^2 \sin^2 \frac{1}{2}\delta}{(1 - b^2)^2 + 4b^2 \sin^2 \frac{1}{2}\delta},$$

where $\delta$ is the phase retardation and is given by the equation

$$\delta = \frac{2\pi}{\lambda} 2e \cos r,$$

and $\lambda$ is the wave length in the material of the film. The intensity will be zero when $\delta = 2n\pi$—that is, when

$$2e \cos r = n\lambda.$$

Hence when $2e \cos r$ is any even number of half wave lengths, there is no light of wave length $\lambda$ reflected from the plate. The light reflected from the upper face is entirely destroyed by that reflected from the lower, and the plate appears perfectly black.

Again, the second form of the value of I shows at once that it is greatest when $\delta = \pi$, or any odd multiple of $\pi$, for in this case $\sin \frac{1}{2}\delta = 1$ and dividing above and below by $\sin^2 \frac{1}{2}\delta$ it appears at once that the denominator is at its least value. The whole fraction is therefore at its maximum value.

It follows then that if

$$2e \cos r = (2n+1)\frac{\lambda}{2}$$

the two streams of light are in the same phase, and the plate appears brilliantly illuminated.

The maximum illumination is measured by

$$\frac{4a^2b^2}{(1+b^2)^2}.$$

This investigation of course applies only to light of a definite wave length. If solar light is used some of these waves will attain their maximum value while others are at their minimum and absent altogether, and the resultant light from the plate will be a mixture of colours which will vary with the thickness of the plate and the angle of incidence. For example, if $\delta$ and $i$ are such that the red light is absent from the emergent beam, then the film will appear of a bluish-green hue. By altering the angle of incidence the light of any particular wave length may be extinguished, and the colour of the film will vary accordingly.

**114. The Transmitted System.**—So far we have taken no account of the light which passes completely through the film, and this we should *a priori* surmise to be complementary to that which is reflected, and therefore the appearance of the film as seen by the transmitted light should be exactly complementary to the appearance presented under the same condition by the reflected light. For the whole light incident on the plate is divided into two portions, one of which returns from the first face, while the other emerges from the second. That this surmise is supported by theory will be readily seen. Consider all the light which emerges from the plate at $C_1$. The first ray is $A_1B_1C_1$. Its amplitude is $acf$, and if its phase at $C_1$ is $\phi$, the equation of its vibration on leaving $C_1$ is

$$y = acf \sin \phi.$$

The next ray which we must take account of is $A_2B_2C_2B_1C_1$, which after incidence at $B_2$ is reflected at $C_2$ and $B_1$, and finally comes to $C_1$. This ray reaches $D_1$ (the foot of the perpendicular from $B_1$ on $B_2C_2$) at the instant the ray $A_1B_1$ reaches $B_1$, consequently the difference of path is $D_1C_2 + C_2B_1$, which is as before, $2e \cos r$. The difference of phase at $C_1$ corresponding to this[1] will be $2\pi/\lambda \,.\, 2e \cos r$, which we have already denoted by $\delta$. The equation of the vibration is therefore

$$y_1 = acfe^2 \sin (\phi + \delta).$$

Similarly the equation of the vibration which comes from $B_3$ along the path $B_3C_3B_2C_2B_1C_1$ will be

$$y_2 = acfe^4 \sin (\phi + 2\delta),$$

and so on.

Summing the series as before, and remembering the relations $b = -e$ and $cf = 1 - b^2$, we find for the square of the amplitude of the resultant vibration

$$\mathrm{I}' = \frac{a^2(1-b^2)^2}{1 - 2b^2 \cos \delta + b^4}.$$

The intensity of the transmitted light will therefore be a maximum if $\delta = 2n\pi$, for then $\cos \delta = +1$ and the denominator of the expression for I′ is least. This greatest value of I′ is simply $a^2$. The maximum value then of the transmitted light is equal to that of the incident, and takes place when $2e \cos r = n\lambda$, but in this case we have found that the intensity of the reflected light is zero. The transmitted light then is a maximum when the reflected light is zero.

Similarly the transmitted light is a minimum when the reflected light is greatest, for when $\cos \delta = -1$, the denominator of the expression for I′ will be greatest. The least value of I′ is then when $\delta = (2n+1)\pi$, or $2e \cos r = (2n+1)\frac{1}{2}\lambda$, and

$$\text{min. } \mathrm{I}' = \frac{a^2(1-b^2)^2}{(1+b^2)^2}.$$

The transmitted light is then never zero, but is always such that when

[1] $\lambda$ is here the wave length in the plate.

added to the reflected light their sum will be equal to the incident light, for we have always

$$\mathrm{I}+\mathrm{I}'=\frac{4a^2b^2\sin^2\frac{1}{2}\delta+a^2(1-b^2)^2}{1-2b^2\cos\delta+b^4}=a^2.$$

The transmitted and reflected lights are therefore always complementary, or the sum of the lights which come from the two faces of the plate will exactly make up the incident light. The difference between the maximum and minimum intensities is the same in both cases—that is, the total variation is the same but the percentage variation is much greater in the reflected system, and the phenomena are consequently much better marked.

### *Example*

Show that if a change of phase is introduced by the act of reflection or refraction, the equations

$$b+c=0, \quad cf=1-b^2$$

still hold, and that the sum of the accelerations of phase at the two reflections is equal to the sum of the accelerations at the two refractions, and the accelerations at the two refractions are equal to each other (Stokes, *Math. and Phys. Papers*, vol. ii. p. 93).

[If the incident vibration be $a\sin\theta$, the reflected and refracted $ab\sin(\theta+\beta)$ and $ac\sin(\theta+\gamma)$ respectively, then on reversal the signs of $\beta$ and $\gamma$ must be changed, so that we have

$$b\sin(\theta-\beta+\gamma)+c\sin(\theta-\gamma+\epsilon)=0,$$

and
$$cf\sin(\theta-\gamma+\phi)=(1-b^2)\sin\theta.$$

Each of these equations must hold for general values of $\theta$, so that the angles added to it in the two terms of these equations must either be equal or differ by some multiple of $\pi$. We may therefore take

$$\phi=\gamma, \quad \text{and therefore} \quad cf=1-b^2,$$

and
$$\beta+\epsilon=2\gamma \quad ,, \qquad ,, \qquad b+c=0.]$$

## SECTION II.—THE COLOURED RINGS OF THIN PLATES

**115. Newton's Rings.**—It has been already mentioned that when two pieces of ordinary plane glass are pressed together the thin film of air enclosed between them generally exhibits a series of highly coloured fringes running in curves round the point of nearest approach of the glasses. When one of the pieces of glass has a spherical surface while the other is plane, as shown in Fig. 96, the layers of equal thickness in the film form a system of concentric circles around the point of nearest approach, and, when viewed in ordinary daylight, a system of highly coloured rings are seen encircling the central spot. These appearances are known as Newton's rings,[1] and they form one of the most beautiful and easily produced examples of interference.

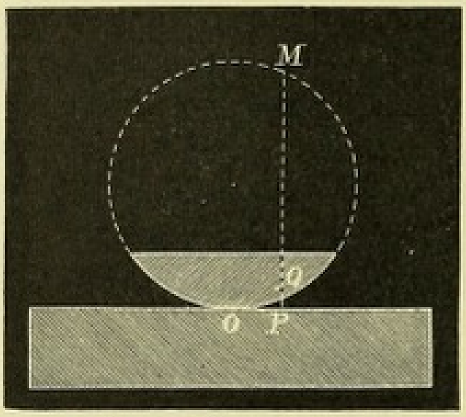


Fig. 96.—Newton's Rings.

The laws according to which these rings are formed are very easily deduced by remarking that the thickness of the film varies approximately as the square of the distance from the point of contact. Thus if OQM be the spherical surface (of radius R) of which the lens is a part, and O its point of contact with the plate of glass, then

$$OP^2 = PQ \times PM = 2Re,$$

since PM is very approximately equal to 2R, the diameter of the sphere, and $PQ = e$ the thickness of the film at P. Now the film consists of circular rings of uniform thickness. Thus at all points of the circle of radius OP around the point of contact the thickness of the film will be the same, and equal to PQ. Denoting the corresponding radius OP by $\rho$ we have the general relation

$$\rho^2 = 2Re,$$

[1] Newton, in studying the formation of these rings, "took two object-glasses, the one a plane-convex for a fourteen-foot telescope, and the other a large double convex for one of about fifty foot; and upon this laying the other with its plane side downwards I pressed them slowly together to make the colours successively emerge in the middle of the circles" (*Opticks*, book ii. p. 172).

consequently if $e$ satisfies the equation

$$2e \cos r = n\tfrac{1}{2}\lambda$$

the ring will appear bright or dark according as $n$ is odd or even.

The radii of the bright rings are therefore given by

$$\rho = \sqrt{R \sec r \,.\, (2n+1)\tfrac{1}{2}\lambda},$$

and the radii of the dark rings by

$$\rho = \sqrt{R \sec r \,.\, n\lambda}$$

when $n$ is any whole number, and $\lambda$ is the wave length in the film.

If $n=0$, then $\rho=0$ and the centre of the system is dark, as we should have expected, since we have supposed that the thickness of the film is zero at this point. The radii of the successive bright and dark rings are proportional to the square roots of the consecutive numbers, the bright rings corresponding to the odd and the dark rings to the even number.

Since the thickness of the film at any point varies as the square of the distance from the centre, it follows that the thicknesses which correspond to the successive rings are proportional to the natural numbers, at the dark rings to the even numbers, and at the bright rings to the odd.

These laws were arrived at with great accuracy by Newton[1] himself, but he did not stop here. He found that in his experiments the absolute thicknesses of the film corresponding to the dark rings were $\frac{2}{178000}$, $\frac{4}{178000}$ in., etc., when the angle of incidence was 4°. From this we find $\lambda = \frac{1}{44500}$ inch approximately, which corresponds to the most luminous part of the spectrum in the neighbourhood of the yellow. These measurements are the first from which the wave lengths of light might have been determined, and Newton made use of them for the purpose of ascertaining the length of a fit, his attention being concentrated on the development of the emission theory.[2]

Wave length determined.

If water instead of air be placed between the glasses the radii of the rings are observed to be much smaller. This then is a proof that light travels slower in water than in air, for here the thickness of

Velocity test.

[1] Newton, *Opticks*, book ii.

[2] "If the rays which paint the colour in the confines of the yellow and orange pass perpendicularly out of any medium into air, the interval of their fits of easy reflection are the $\frac{1}{89000}$th part of an inch. And of the same length are the intervals of their fits of easy transmission" (Newton, *Opticks*, book ii. part iii. prop. 18).

For the thicknesses at the dark rings were $\frac{2}{178000}$, $\frac{4}{178000}$, etc. This obviously corresponds to half a wave length, so that we have for yellow light the first determination of $\lambda = \frac{1}{44500}$ in.

water required to retard one component on the other by a definite amount is less than the thickness of air, which produces the same result. The formula which determines the radii of the rings also points to the same conclusion, for the wave length in any medium is proportional to the velocity, and we infer that the radius of any ring varies approximately as the square root of the velocity in the film when the incidence is nearly normal. Thus, from the contraction exhibited by the rings when water or any other fluid replaces air between the lens and plate, it is possible to compare the velocities of light in these media.

When violet light is used the rings are smaller[1] than with red light, and the theory points out that the ratio of the radii are as the square roots of the wave lengths; hence we have again arrived at the conclusion that the violet waves are shorter than the red, and we can again not only compare their magnitudes, but absolutely determine their lengths in any given substance.

When ordinary solar light is used a series of iris-coloured rings are exhibited, violet at the inner and red at the outer edge. The order of succession of the colours laid down by Newton[2] from the centre outwards for the successive rings was: (1) black, blue, white, yellow, red; (2) violet, blue, green, yellow, red; (3) purple, blue, green, yellow, red; (4) green, red; (5) greenish-blue, red; (6) greenish-blue, pale red; (7) greenish-blue, reddish-white. This list is generally referred to as "Newton's Scale of Colours."

Thus we read of the red or blue of the "third order," meaning thereby that red or blue which is seen in the third rainbow-coloured ring which encircles the central dark spot.

With white light the alternations are few, for the coloured rings soon become superposed and overlapped, so as to obliterate all traces of interference and colour, and the rings fade gradually into uniform illumination.

**116. The Transmitted Rings.**—The rings we have spoken of so far are produced by the interference of the streams of light reflected from the two surfaces of the thin film. It is obvious that the light transmitted through the film should also exhibit interference phenomena, but of a complementary character, the maxima and minima of one system corresponding to the minima and maxima of the other. Thus when the film is looked at from the other side a system of

[1] This may be observed (following Newton) by illuminating the glasses with light from different parts of the spectrum, or more simply, by looking at the rings, formed by ordinary light, through differently coloured glasses.

[2] *Opticks*, book ii. obs. 4; see also table, p. 228.

rings is observed, complementary in character to those observed by reflection, and consequently encircling a white centre. These rings are much paler than the reflected system, for on account of the great difference in the intensities of the interfering pencils,[1] there are no points of absolute darkness, so that the bright rings do not stand out so prominently as in the reflected system.[2]

If the two ring systems are viewed at once it follows that uniform illumination should be the result. This has been verified by Arago.[3] Placing a glass plate and lens in contact in a vertical position over a horizontal sheet of uniformly illuminated white paper, an eye situated at E (Fig. 97) will receive light from B by reflection at C, and also light from A by transmission. Both the reflected and transmitted systems are in the field of view, but being exactly complementary, the result is uniform illumination.

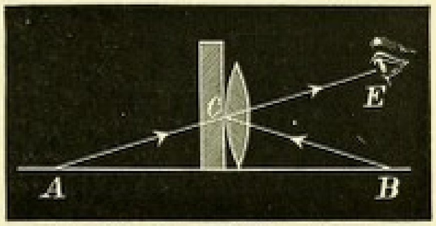


Fig. 97.—Arago's Experiment.

Specimens of ancient glass sometimes show transmitted colours of great brilliancy. Brewster's explanation is that owing to superficial decomposition of the glass we have here to deal with a series of thin plates of nearly equal thicknesses. With such a series the transmitted colours should be much purer, and the reflected much brighter than is usual with a single plate.[4]

**117. Rings with a White Centre.**—The system of rings produced by the transmitted light as we have observed starts from a white centre, but the feature of the reflected system is that their

[1] When light is incident perpendicularly on glass about 4 per cent is reflected; the intensity of the first reflected beam is therefore about $\frac{1}{25}$ of the incident beam. The rest enters the glass and loses 4 per cent again by reflection at the second surface, so that $(\cdot 96)^2$, or ·9216 of the original light passes through. The light transmitted after being twice reflected inside the plate is $I(\cdot 96)^2 (\frac{1}{25})^2$, or about $\frac{1}{6}$ per cent. That four times reflected inside will only be $\frac{1}{625}$ of this, and so on. The difficulty then is to understand how such a small quantity of light when superposed on the strong direct beam should produce any perceptible rings at all. However the intensity of the weaker beam is $(\frac{1}{25})^2$ if the intensity of the stronger is unity, hence the amplitude of the weaker is $\frac{1}{25}$, and the maximum and minimum intensities in the transmitted beam will be

$$(1 \pm \tfrac{1}{25})^2 = 1 \pm \tfrac{2}{25}\ q.p.,$$

so that the difference is as much as $\frac{4}{25}$ of the stronger beam. (For the complete calculation see Art. 113.)

[2] Noticed by Newton, *Opticks*, book ii. obs. 9.

[3] Arago, *Œuvres complètes*, tom. x. p. 16 (note).

[4] The analytical investigations of Stokes for a pile of plates (*Proc. Roy. Soc.* vol. xi. p. 545, 1860) may be applied to this question.

central spot is black. The theory accounts for the black spot by showing that the two interfering streams of light are reflected under different conditions, one in passing from dense to rare, and the other in passing from rare to dense, the result being that a difference of phase of half a period is introduced between the two beams. Let us now consider the case of two plates enclosing a film of a refractive index intermediate between those of the plates themselves. Thus suppose the film to be denser (more refracting) than the first plate, and less refracting than the second plate, then the reflection at the first surface of the film will take place when the light is passing into a more refracting medium, and the same will be the case at the second surface also. The light is therefore reflected under similar circumstances at both faces, and no difference of phase is introduced. It follows, then, that in the system of rings formed under these circumstances the central spot should be white.

Young verified this anticipation of the theory by enclosing oil of sassafras between two lenses, one of which was of flint glass, and the other of crown glass. By this experiment Young justified the hypothesis of the loss of half an undulation.

The oil of sassafras may be replaced by a mixture of essence of cloves and essence of laurels. If the film be of higher index than the object-glasses between which it lies, the centre of the rings should still be black. Arago verified this by using oil of cassia, of which the index is superior to that of flint glass.

When the third medium differs from the first, the theory of thin plates becomes more complicated. In one case, however, no colours at all should be exhibited, viz. when the film is backed by a perfect reflector, such as polished silver covered with a film of gelatine. In this case the waves are reflected *in toto*, so that the reflected and transmitted systems become superposed.

**118. Conditions for Large and Bright Rings.**—The formula of Art. 115 shows that the diameter of the $n$th ring increases with the radius of curvature of the lens—that is, with the tenuity of the film. Hence, in order to obtain wide rings, a lens of very small curvature should be employed, but with a given piece of apparatus there is still another factor to be considered in estimating the magnitude of the rings, viz. the angle of refraction into the film. The diameters of the rings depend on the secant of this angle, and they therefore increase with it. With a simple piece of apparatus, such as that shown in Fig. 96, when the angle of incidence is increased, there is great loss of light by reflection at the first or upper surface of the glass, and only a small fraction of the incident beam reaches the film, so that a large angle of incidence is detrimental to brightness, and the rings obtained will be very faint. This difficulty may be avoided by using a prism and a lens, as shown in Fig. 98, instead of the plate and lens employed by Newton. One face of the prism is placed on the curved surface of the lens so that light falling nearly perpendicularly on one of the other faces will enter the prism in large quantity, and in such a direction that the angle of refraction into the film is also large. Viewed through the third face of the prism the rings obtained in this manner are both bright and large. The same result is obtained with a glass plate and a prism having one face polished into a spherical form of small curvature.

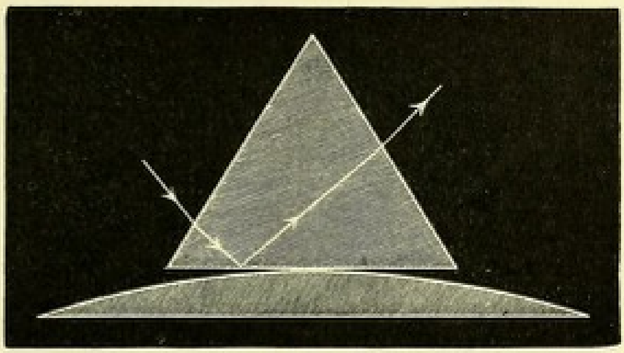


Fig. 98.

The chief peculiarity of the rings obtained in this manner is not so much the brilliancy, or the diminished influence of the thickness of the film, as the more or less perfect achromatism produced by dispersion, especially in the neighbourhood of total reflection, so that the bright rings are nearly white instead of being highly coloured. This happens because the more refrangible rays are more deviated by the prism, and therefore enter the film at a greater angle, so that for a given thickness of film they have a smaller path retardation, as explained in Art. 110. This means that the dispersion increases the diameters of the rings corresponding to the more refrangible rays, and when the diameters of the rings of a given order are the same for all wave lengths there is perfect achromatism, and the rings are black and white. This compensating effect of dispersion diminishes the confusion which arises from overlapping, and increases the number of visible rings. As the incidence augments the violet rings may become larger than the red, and a reversal of colour occurs. This takes place near total reflection, which happens first for the violet light, and is attended by a rapid change in the value of cos $r$ (see further, Art. 120).

Achromatising effect of dispersion.

**119. Examination of Newton's Rings through a Prism.**—When Newton's rings are examined through a prism some remarkable phenomena are exhibited. They are described in his twenty-fourth observation, *Opticks* (book ii.):—

"When the two object-glasses are laid upon one another so as to make the rings of the colours appear, though with my naked eye I could not discern above eight or nine of those rings, yet by viewing them through a prism I could see a far greater multitude, insomuch that I could number more than forty . . . and I believe that the experiment may be improved to the discovery of far greater numbers. . . . But it was on but one side of these rings, namely, that towards which the refraction was made, which by the refraction was rendered distinct, and the other side became more confused than when viewed with the naked eye. . . .

"The arcs where they seem distinctest were only black and white successively, without any other colours intermixed.

"I have sometimes so laid one object-glass upon the other that to the naked eye they have all over seemed uniformly white without the least appearance of any of the coloured rings; and yet by viewing them through a prism great multitudes of those rings have discovered themselves. And in like manner plates of Muscovy glass and bubbles of glass blown at a lamp furnace, which were not so thin as to exhibit a great variety of them, ranged irregularly up and down in the form of waves. And so bubbles of water, before they began to exhibit their colours to the naked eye of a bystander, have appeared through a prism, girded about with many parallel and horizontal rings; to produce which effect it was necessary to hold the prism parallel, or very nearly parallel, to the horizon, and to dispose it so that the rays might be refracted upwards."

Newton attributes these "odd circumstances" to the dispersing power of the prism. The blue being more refracted than the red, it is possible that the $n$th blue ring may be so displaced relatively to the $n$th red ring that, at part of the circumference, the displacement may compensate for the difference of diameters. A white strip may thus be formed in a situation where without the prism the mixture of colours would be complete, so far as could be judged by the eye.

A simple case is that in which the thin film is a wedge bounded by plane surfaces inclined at a small angle. If the edge of the prism is parallel to the intersection of the faces of the plate, by drawing back the prism it will be possible to adjust the effective dispersing power so as to bring the $n$th bars to coincide for any two assigned colours, and therefore approximately for the entire spectrum. The formation of these achromatic bands depends upon the same principles as the fictitious shifting of the centre of a system of Fresnel's bands when viewed through a prism (Art. 104).

**120. Herschel's Fringes.**—A very simple and effective method of obtaining coloured fringes by reflection from a thin plate of air was first pointed out by Sir William Herschel.[1] On a perfectly plane piece of glass or a metallic mirror, before an open window, place an equilateral prism. The light falling upon the exposed face of the prism is reflected at the base and emerges from the other face (Fig. 99). To an observer looking in through the latter face the field appears divided into two parts, one brightly illuminated, which arises from the occurrence of total reflection at the base of the prism, and the other comparatively dark, the light there being partly transmitted.

[1] Sir William Herschel, *Phil. Trans.* 1809, p. 274.

The line of separation of the two parts would be circular to an eye situated in the prism, but the apparent shape of the curve seen by the eye outside is the distorted form of a circle seen by refraction through the prism. Since total reflection occurs at slightly different incidences for the different colours, it follows that the curve of separation will be iris-coloured. Inside this coloured band, and running parallel to it, we have, in addition, a system of beautifully coloured fringes, the breadth and number of which vary with every change of pressure.

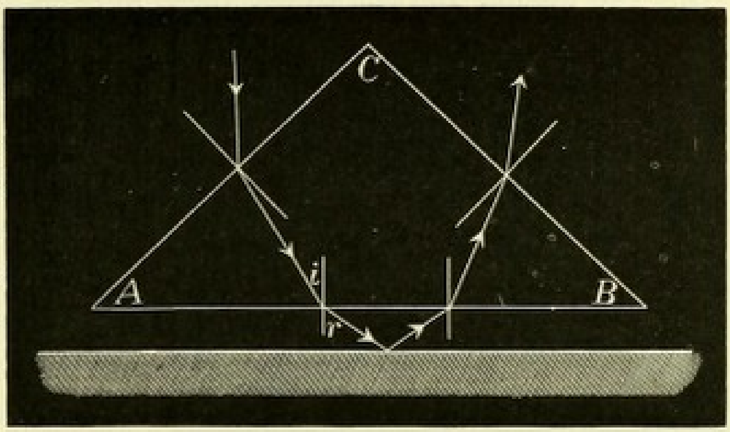


Fig. 99.

These bands do not require for their formation a perfect polish in the lower surface nor extreme thinness in the air film, for they may be seen very well when the prism is separated from the lower plate by the thickness of thin tissue paper or a fibre of cotton wool.

That excessive thinness of the air film is not necessary to the production of tolerably broad fringes when the angle of incidence approaches the angle of total reflection may be inferred at once from the expression $2e \cos r$, which gives the path retardation of the interfering pencils in the case of a film of thickness $e$. Near the critical angle $r$ is nearly $90^\circ$, and $\cos r$ is very small, so that the retardation may be small even with a sensible thickness of film.[1]

When the prism and plate combined are held up to the light a transmitted iris is seen, lined with a similar system of fringes, on looking through the plate and the base of the prism.

The experiment may also be conducted by merely looking through two prisms placed in contact (Fig. 100). In this form it was repeated by H. F. Talbot.[2] If the prisms be equal, isosceles, and right-angled, then when placed with their hypothenuses in contact their ends will form squares. Looking through the combined prisms at the sky, a system of bands is seen, and looking at the interface so as to see the

[1] These fringes may also be obtained very conveniently from a film of air enclosed between two plane glass plates which are separated by two fragments of the same thread of platinum wire (about $\frac{1}{10}$ mm. in diameter), cemented around their edges, and plunged vertically in water contained in a rectangular glass vessel in the manner described in Art. 83.

[2] H. F. Talbot, *Phil. Mag.* 1836, p. 401.

light reflected there, another system is observed, the latter being complementary to the former.

Fox Talbot describes a modification of the experiment as follows :— "But the beauty of the appearances may be surprisingly increased by transporting the apparatus into a dark chamber and suffering a pencil of the brightest solar light to pass through the prism, or to be reflected from the face AC. If then a sheet of white paper be held up, at any distance from the prism, the coloured bands are depicted upon it with the greatest vivacity and distinctness. The transmitted bands have altogether a different *character* from the reflected ones, so that it is impossible to mistake one for the other, even without reference to the path of the ray.

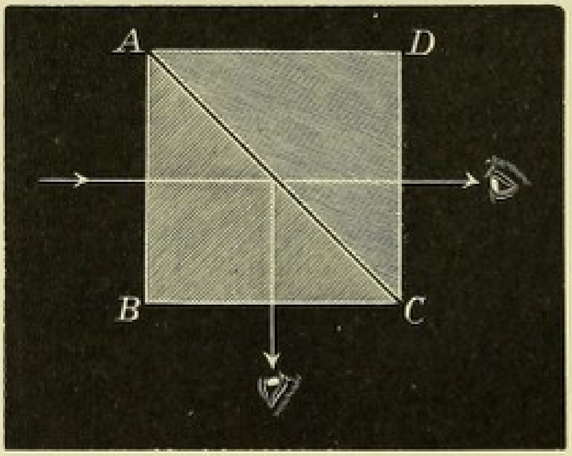



Fig. 100.

"The coloured bands are not, as has been supposed, *isochromatic* lines. The deviation is sometimes very marked, so that a band in the course of its progress acquires very different tints from those which it possessed originally. This fact may be considered of some importance with respect to the theory. It takes place when the prisms are in close contact and the bands few in number. But the following is still more deserving of attention. When the contact of the prisms is diminished by interposing a hair between them (still pressing them together), the coloured bands depicted upon the paper become more numerous, narrow, and crowded. Frequently they alternate a great number of times with two complementary colours. This appeared to me so remarkable that I repeated the experiment with additional care. The radiant point of solar light was made smaller by transmitting the ray through a lens of short focus, and the position of the combined prisms was slowly altered by turning them round their centre. The appearance of the bands on the paper was all the time carefully noted. I soon found a position of the prisms in which the remarkable phenomenon occurred of a complete compensation of colour—that is to say, that the bands were black and white. At the same time they were become exceedingly narrow and numerous. . . . They resembled more than anything else the closely-ruled parallel lines by which *shadows* are produced in some kinds of engraving, and which are often employed in maps to represent the sea.

"Now it requires in ordinary circumstances the employment of *very homogeneous* light, in order to produce bands anything like these in number and distinctness. In the present instance, on the contrary, common solar light was employed. . . . These bands are best seen in the light *reflected* from the face AC."

The achromatism here referred to is accounted for by the dispersion which takes place in the glass, and which becomes highly effective when the light falls upon the film at an angle nearly equal to the critical angle, so that the angle of refraction into the film is nearly 90°, as explained in Art. 110.

In the first place, it may be remarked that when the prism is isosceles, so that the angle A is equal to the angle B, then a ray entering the first face AC (Fig. 99) at any angle will leave the second face BC at the same angle, and consequently a beam of white light entering the face AC will leave the face BC without dispersion. In other words, the dispersion produced in the angle A will be compensated in the angle B. The function of the prism is therefore to vary the angle of refraction $r$ into the film in such a way that $\cos r/\lambda$ is approximately the same for all the colours. When A is not equal to B there will be a further displacement of the fringes such as would be produced by regarding them through a prism of angle A—B, as described in Arts. 105 and 119.

Let the angle of incidence on the first face of the prism be $i_1$ and the corresponding angle of refraction $r_1$. Then if $i$ and $r$ be the angles of incidence and refraction for the film, we have

$$\sin i_1 = \mu \sin r_1 \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (1)$$

$$\sin r = \mu \sin i \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (2)$$

$$i + r_1 = \text{A} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (3)$$

where $\mu$ is the refractive index of the prism, and the film is supposed to be air. When the film is other than air, equation (2) can be modified accordingly. These equations combined with the achromatic condition

$$\frac{\cos r}{\lambda} = \text{const.} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (4)$$

determine the angle of incidence at which achromatism takes place, when the law of dispersion in the glass is known.

Thus in Talbot's form of the experiment a parallel beam of white light was used, so that $i_1$ is the same for all wave lengths; hence by differentiating the foregoing equations we obtain

$$\mu \cos r_1 dr_1 + \sin r_1 d\mu = 0 \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (1')$$

$$\mu \cos i di + \sin i d\mu = \cos r dr \quad . \quad . \quad . \quad . \quad . \quad . \quad (2')$$

$$di + dr_1 = 0 \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (3')$$

$$\lambda \sin r dr + \cos r d\lambda = 0 \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (4')$$

Combining these equations we obtain at once M. Mascart's relation [1]

$$\cot^2 r = -\frac{\sin A}{\sin r \cos r_1}\lambda\frac{d\mu}{d\lambda},$$

which may be written in the equivalent form [2]

$$\cot^2 r = -\frac{\sin A}{\sin i \cos r_1}\frac{\lambda}{\mu}\frac{d\mu}{d\lambda}.$$

The foregoing applies to Talbot's form of the experiment in which the angle of incidence $i_1$ is constant, and all the light of a given colour is refracted at the same angle, so that the bands arise from variations of the thickness of the film, such as when a hair is placed between the prisms to produce a wedge-shaped film.

In Herschel's form of the experiment the prism is placed before the open sky, so that the angle of incidence is variable and light of a given wave length is not all refracted at the same angle. These bands are broad and richly coloured, and they are produced near the limit of total reflection by the variation of the angle of incidence when the thickness of the film is constant. Of course bands of an intermediate character are produced when the angle of incidence and the thickness of the film both vary.

The theoretical condition for constant thickness is better satisfied, if, after Mascart, we place the layer of air in the focus of a small radiant point (electric arc). In this case the area concerned may be so small that the thickness in operation can scarcely vary, and the ideal Herschel's bands are seen depicted on a screen held in the path of the reflected light. It will, of course, be understood that bands may be observed of an intermediate character in the formation of which both thickness and incidence vary. Herschel's relate to one particular case—that of constant thickness; Talbot's to the other especially simple case of constant angle of incidence.

From the present point of view there is one very important distinction, because one is achromatic and the other is not. To understand this, we follow Herschel's bands in greater detail.

In the same notation as before

$$\delta = 2e\cos r = n\lambda,$$

and the question to be investigated is the relation of $i_1$ to $n$. The band of zero order ($n=0$) occurs when $r=90°$, that is at the critical angle. For two successive bands we have

$$n\lambda = 2e\cos r, \quad (n+1)\lambda = 2e\cos(r+dr),$$

[1] M. Mascart, *Traité d'Optique*, tom. i. p. 449.

[2] Lord Rayleigh, *Phil. Mag.* vol. xxviii. p. 196, 1889.

therefore

$$\lambda = -2e \sin r dr.$$

Also

$$\cos i_1 di_1 = \mu \cos r_1 dr_1,$$

so using previous results we find

$$di_1 = \frac{\mu \cos r_1}{\cos i_1} \cdot \frac{-\cos r}{\mu \cos i} \cdot \frac{-\lambda}{2e \sin r} = \frac{n\lambda^2}{4e^2} \frac{\cos r_1}{\cos i_1 \cos i \sin r}.$$

Near total reflection $\sin r = 1$, (*q.p.*) and the factors $\cos r_1, \cos i_1, \cos i$ vary but slowly with the order of the band and also with the wave length. Hence the width of the band is approximately proportional to the order, the square of the wave length, and the inverse square of the thickness.

When the light is white, the centre of the system will be where there is coincidence of bands of the order $n$ in spite of the variation of $\lambda$. About the achromatic centre thus determined the visible bands will be grouped. At the central band $n$ is the same for the various colours, consequently the widths of the various systems are at this place approximately proportional to $\lambda^2$. Hence these bands are less achromatic than ordinary bands or Newton's rings, in which the width is proportional to $\lambda$, and this theoretical conclusion is in harmony with observation (see Lord Rayleigh's paper, *Phil. Mag.* Sept. 1889).

### *Newton's Observations on the Coloured Rings of Thin Plates*

*Opticks*, book ii. part i. obs. 4: "I took two object-glasses, the one a plane-convex for a fourteen-foot telescope, and the other a large double convex for one of about fifty foot; and upon this laying the other with its plane side downwards I pressed them slowly together, to make the colours successively emerge in the middle of the circles, and then slowly lifted the upper glass from the lower to make them successively vanish again in the same place. The colour, which by pressing the glasses together emerged last in the middle of the other colours, would upon its first appearance look like a circle of a colour almost uniform from its circumference to its centre, and by compressing the glasses still more, grew continually broader till a new colour emerged at its centre, and thereby it became a ring encompassing that new colour. And by compressing the glasses still more the diameter of this ring would increase, and the breadth of its orbit or perimeter decrease until a new colour emerged in the centre of the last; and so on until a third, a fourth, a fifth, and other following new colours successively emerged there, and became rings encompassing the innermost colour, the last of which was the black spot. And, on the contrary, by lifting up the upper glass from the lower, the diameter of the rings would decrease, and the breadth of their orbit increase, until their colours reached successively to the centre; and then by being of a considerable breadth, I could more easily discern and distinguish their species than before. And by this means I observed their succession and quantity to be as followeth.

"Next to the pellucid central spot made by the contact of the glasses succeeded blue, white, yellow, and red. The blue was so little in quantity that I could not discern it in the circles made by the prism, nor could I well distinguish any violet in it, but the yellow and red were pretty copious, and seemed about as much in extent as the white, and four or five times more than the blue. The next circuit in order of colours immediately encompassing these were violet, blue-green, yellow, and red ; and these were all of them copious and vivid, excepting the green, which was very little in quantity, and seemed much more faint and dilute than the other colours. Of the other four the violet was the least in extent, and the blue less than the yellow and red. The third circuit or order was purple, blue, green, yellow, and red ; in which the purple seemed more reddish than the violet in the former circuit, and the green was much more conspicuous, being as brisk and copious as any of the other colours except the yellow ; but the red began to be a little faded inclining very much to purple. After this succeeded the fourth circuit of green and red. The green was very copious and lively, inclining on the one side to blue and on the other side to yellow. But in this fourth circuit there was neither violet, blue, nor yellow, and the red was very imperfect and dirty. Also the other colours became more and more imperfect and dilute, till after three or four revolutions they ended in perfect whiteness."

"*Obs.* 5.—To determine the interval of the glasses, or thickness of the interjacent air, by which each colour was produced, I measured the diameters of the first six rings at the most lucid part of their orbits, and squaring them, I found their squares to be in the arithmetical progression of the odd numbers, 1, 3, 5, 7, 9, 11. And since one of these glasses was plane and the other spherical, their intervals at those rings must be in the same progression. I measured also the diameters of the dark or faint rings between the more lucid colours, and found their squares to be in the arithmetical progression of the even numbers, 2, 4, 6, 8, 10, 12. And it being very nice and difficult to take these measures exactly ; I repeated them divers times at divers parts of the glasses, that by their agreement I might be confirmed in them."

"*Obs.* 10.—Wetting the object-glasses a little at their edges, the water crept in slowly between them, and the circles thereby became less and the colours more faint ; insomuch that as the water crept along, one half of them at which it first arrived would appear broken off from the other half, and contracted into a less room. By measuring them I found the proportions of their diameters to the diameters of the like circles made by air to be about seven to eight, and consequently the intervals of the glasses at like circles, caused by those two mediums water and air, are as about three to four. Perhaps it may be a general rule that if any other medium more or less dense than water be compressed between the glasses, their intervals at the rings caused thereby will be to their intervals caused by interjacent air, as the sines are which measure the refraction made out of that medium into air."

## SECTION III.—THE COLOURS OF THICK PLATES

**121. Brewster's Bands.**—When a pencil of light falls in succession upon two transparent plates which are not very thin,[1] some of the many portions into which it is divided by partial reflections at their bounding surfaces are frequently in a condition to interfere and produce coloured bands. These fringes were observed by Sir D. Brewster [2] in 1815. The apparatus employed in the experiment consisted of a straight tube (Fig. 101) blackened on the inside and closed at one end by a disc containing a small aperture O. Two uniform [3] glass plates of equal thickness were placed near each other at the other end of the tube, one of them being at right angles to the axis of the tube, and the other inclined to it at a very small angle. This angle could be varied by means of a micrometer screw.

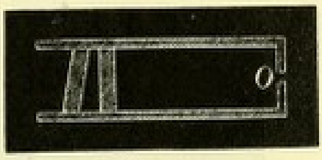


Fig. 101.

In order to understand the formation of these fringes it is only necessary to notice that when the aperture O is illuminated and viewed through the plates the light which reaches the eye consists of many distinct components arising from successive reflections within and between the plates, and a corresponding series of images of the source is consequently presented.

This series of images may be divided into a system of groups. Those of the first group correspond to light that has traversed the space between the plates once, so that they are formed by light that has not been reflected from one plate to the other, but which may have suffered reflection within either plate. The first image of this

[1] A thick plate in optics means one of thickness large compared with the length of a wave of light.

[2] Brewster, *Edinb. Trans.* vol. vii. p. 435, 1815.

[3] A plate of perfectly uniform thickness cannot be procured in practice, and it is difficult to obtain plates sufficiently uniform to give good fringes by this method. A plate of glass generally has its faces inclined to each other so as to form a prism of very small angle. The lines of constant thickness in such a plate are approximately rectilinear and parallel, and can be seen when the plate is viewed by reflection in monochromatic light. The interference fringes follow the lines of constant thickness, and if the plate be cut in two along a line perpendicular to the direction of the fringes, the two parts when superposed by folding them round the line of section will readily exhibit Brewster's bands, for in this case the corresponding rays of the interfering pencils traverse the plates at places of equal thickness.

group may be called the principal image, and it is formed by light that has been directly transmitted through the plates, such as AA in Fig. 102. Behind this there is an image formed by light that has suffered two internal reflections in one of the plates, such as the rays BB and CC, and so on for multiple internal reflections. Now the image formed by BB will coincide with that formed by CC when the plates are parallel and of equal thickness; but when the plates are slightly inclined the thickness traversed by the light in one of them will differ by a small amount from that traversed in the other, and there will be a small relative path retardation introduced between the pencils BB and CC, so that interference bands will be produced. Regarding the image as a source of light we may say, then, that the image formed by BB interferes with the image formed by CC, and

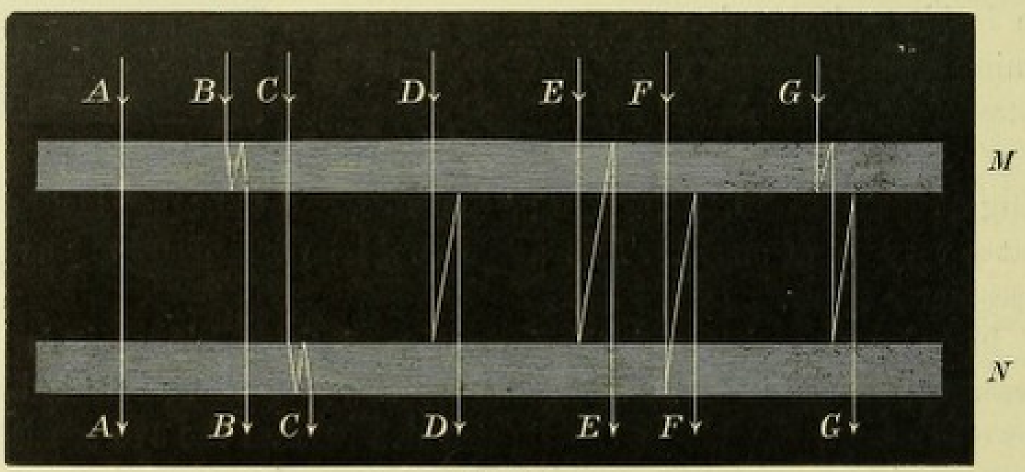



Fig. 102.

produces an image crossed by fringes. The image formed by AA presents no fringes, for there is no other pencil of light of approximately the same path. The other images of this group are also crossed by bands — for example, the light that has suffered four reflections in the plate M interferes with that which is four times reflected within the plate N, etc. The second group consists of light that has traversed the space between the plates three times. The first image of this group is formed by light that has traversed each plate only once, such as the ray DD, and, like the first image of the first group, it presents no fringes. A little behind this, however, comes the image formed by EE, and this is interfered with by FF, so that fringes are presented. Other images are formed by rays, such as GG, that have suffered four or more reflections, and they also present bands which are explained in the same manner. The third group is formed by light that has traversed the space between the plates five times, and so on.

The relative retardation of the pencils forming the interference bands on the second image of any group is easily expressed in terms of the angles of refraction into the plates. For if $e$ be the common thickness of the plates it is clear that the ray EE suffers a relative retardation $2e \cos r$ in the plate M, while FF suffers a relative retardation $2e \cos r'$ in N. The two transmitted pencils have consequently a relative path retardation of

$$\delta = 2e(\cos r - \cos r').$$

This vanishes when $r = r'$—that is, when the light is incident on the two plates at the same angle, or is parallel to the plane bisecting the obtuse angle between the plates. The light incident in this direction consequently determines the central fringe of the system. Brewster describes the experiment as follows :—

"In order to observe the phenomenon to the greatest advantage, let the light of a circular image subtending an angle of 1° or 2° be incident perpendicularly, or nearly so, upon two plates of parallel glass placed at a distance of one-tenth of an inch, and let one of the plates be gently inclined to the other, till one or more of the reflected images be distinctly separated from the bright image formed by transmitted light and received upon the eye placed behind the plates. Under these circumstances the reflected image will be crossed with about 15 or 16 beautiful parallel fringes . . . the direction of the fringes is always parallel to the common section of the four reflecting surfaces.

"All the preceding experiments were made with plates which were cut out of the same piece of glass, and had therefore the same thickness. I now tried plates of different thicknesses, both when ground parallel and when cut from a common plate of glass ; but I could never render the coloured fringes visible, unless when the glass was parallel, and exactly of the same thickness in both plates."

**122. Jamin's Interference Refractometer.** — The interference bands obtained by Brewster's method have been turned to account by M. Jamin[1] in the construction of a very delicate refractometer.

Two plates of parallel glass as nearly as possible of equal thickness (about 1 cm.) are silvered on their backs and supported (on an optical bench or otherwise) so that the distance between them can be altered at will. Their plane faces can be placed in the vertical, which is supposed perpendicular to the plane of the paper in Fig. 103.

The first plate AB is fixed, and the light of the sun or any other source falls upon it. After reflection it is received on the second plate CD, which can be turned round a horizontal axis by means of a screw situated at its back, and round a vertical axis by means of the screw Q. The displacement of this plate round the vertical is measured by the movement of an arm on a graduated arc at V. Part of the light incident on the first face AB is reflected there, and after penetrating the second plate is reflected at its second surface

[1] Jamin, *Ann. de Chim. et de Phys.* third series, tom. lii. p. 163, 1858.

and emerges from the plate. A second part of the light penetrates the first plate, and after reflection at its second surface it is reflected from the first surface of the second plate. Thus of the two beams one is reflected at the front of the first plate and the back of the second, while the other is reflected at the back of the first plate and the front of the second.[1] Hence if the plates be parallel, the two pencils will traverse equal and similar paths, and they will therefore be in the same phase on leaving the second plate; but if the second plate be turned through a small angle the ray which is refracted in it will pursue a path slightly different from that traversed by the ray refracted in the first plate; there will therefore be a difference of

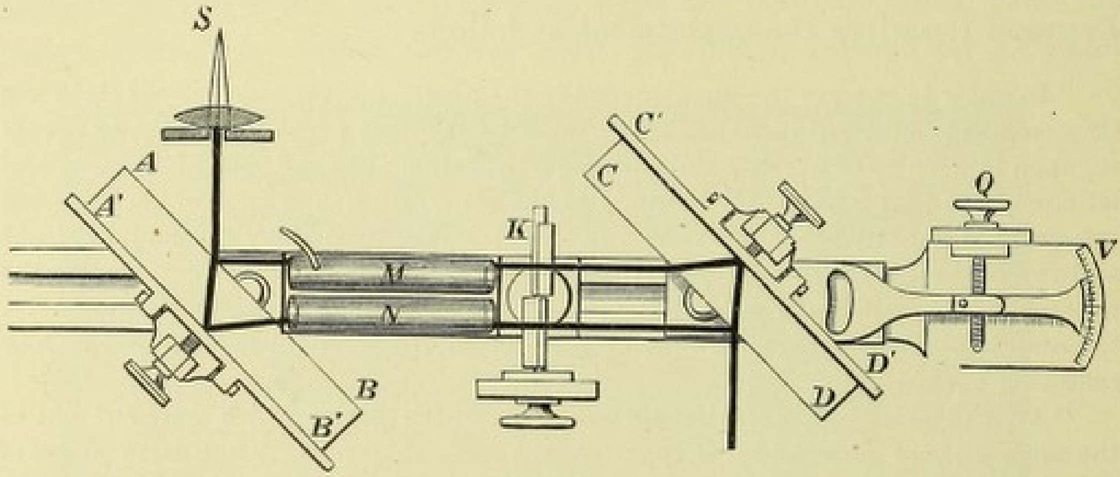



Fig. 103.

phase between the beams as they leave the second plate, and the phenomena of interference will be presented.

If we start with the two plates parallel and the plane of incidence horizontal, then by turning the screw at the back of CD the plates will become inclined to each other and their line of intersection will be horizontal. A system of horizontal fringes will consequently appear in the field, increasing in width as the inclination of the plates is diminished. When the plates are rigorously parallel the

[1] There is of course a whole system of images formed by successive reflections within the plates and at their bounding surfaces. Thus an image is formed by light reflected directly from the face AB and then from CD. Behind this there is the image considered in the text and which is crossed by bands because it consists of two nearly coincident images—namely, that formed by reflection from AB and C'D' combined with that formed by reflection from A'B' and CD. The next image is bright and may be considered the principal image of the system, and is formed by the light reflected from the metallic surfaces A'B' and C'D'. This is followed by other fainter images produced by light that has suffered two or more internal reflections within the plates. These are also crossed by interference bands, and the whole system is situated sensibly on the perpendicular to the plates drawn through the source.

bands should disappear entirely and give place to a uniform illumination. As this stage is approached, however, the bands generally become deformed and form undulating curves indicating want of homogeneity in the glass or imperfections in the planeness of the surfaces.

The central fringe corresponds to zero retardation and, as in Brewster's experiment, is formed by those rays of the interfering pencils which make the same angle with the two plates. These rays are consequently parallel to the plane bisecting the obtuse angle between the plates. The effect of moving the screw Q is to displace the central band vertically and raise or lower the whole fringe system in the field of view.

When a slender pencil of light (from a slit) is used, tubes containing gases or thin plates of different substances may be placed in the paths of the pencils between the plates, and the relative speeds of light in these substances determined by the corresponding displacement of the fringes. The apparatus may thus be used as a refractometer.

The retardation produced by the passage of one of the pencils through a thin plate of any substance, or through a tube of gas, of which it is desired to measure the refractive index, is determined by means of a *compensator* K. Thus if a tube of gas or a thin plate be in the path of one pencil, a thin plate of parallel glass may be introduced across the path of the other, and if this latter be of the proper thickness, there will be no resultant displacement of the fringes. The one will compensate the other. The proper thickness of the parallel glass plate may be adjusted by constructing it of two slender prisms of equal angle, placed on each other with their edges in opposite directions. Placed thus they form a parallel plate, the thickness of which may be varied at will by sliding one prism upon the other. This may be done conveniently by keeping one fixed, and displacing the other by means of a screw. When the instrument is once standardised we know immediately the retardation produced by the substance in the path of the first pencil.

The compensator used by M. Jamin consisted of two glass plates (shown Fig. 104), fixed to a common axis, and inclined to each other at a small constant angle. One pencil passes through one plate, and the other through the other plate. The retardation produced by a plate depends on the angle of incidence, and is least when the ray passes through it normally. When the rays fall on the plates perpendicularly to the plane bisecting the angle between them, they introduce equal retardations, and therefore no displacement of the fringes. In any other position the plates displace the fringes by an amount depending on their thickness and the angles of incidence; so that, by rotating the axis to which they are fixed, any desired displacement of the fringes can be produced, and any previous displacement compensated. The sensitiveness of this compensator can be varied at will by altering the angle between the plates, and the retardation introduced by it may be easily calculated, but it is best to graduate it by experiment, which shows that the relative retardation is very nearly proportional to the angle through which it is turned from the position of zero displacement. M. Jamin by this method has determined the index of refraction of several gases, and compared that of dry and moist air. By the same process he also investigated the effect of compression and change of temperature on the refractive index of water.[1]

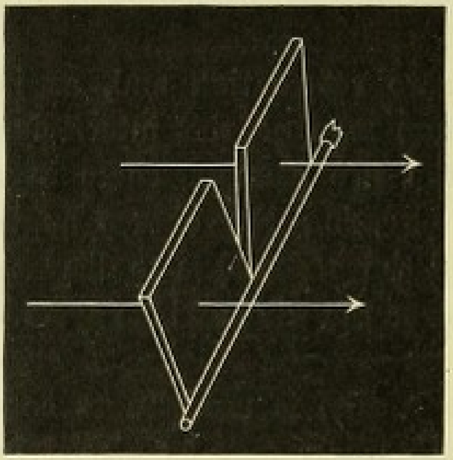


Fig. 104.—The Compensator.

**123. Michelson and Morley's Interferometer.**—An interference refractometer has been described by Professors Michelson and Morley[2] which readily permits of the introduction of any relative retardation between the interfering pencils, and consequently allows of the observation of interference bands corresponding to a large difference of path. This apparatus is the same as that employed in their experiments on the relative motion of the earth and the ether (Art. 315).

Light from a source of light (a sodium flame) S (Fig. 105) falls upon a plane glass plate A inclined to it at any angle (usually 45°). Part of this light is transmitted in the direction AB and part is reflected in the direction AD. Both of these beams are received perpendicularly on plane mirrors C and D, so that they return to the plate A along their original paths and pass in part into the observing telescope T. Now the pencil reflected from D traverses the plate A three times before it reaches T, and in order to compensate for this a similar plate of glass B is introduced into the path of the pencil reflected from C. Hence if AC = AD, and if the plates are parallel, the two pencils will have traversed paths of equal length, but it is to be observed that they have both suffered reflection at the same face of A, one internally and the other externally, and consequently when

[1] Jamin, *Ann. de Chim. et de Phys.* third series, tom. lii. p. 163, 1858; and tom. lxi. p. 385, 1861.

[2] Professors A. A. Michelson and E. W. Morley, *Journal of the Association of Engineering Societies*, May 1888.

the adjustment of the mirrors is exact the whole field will be dark. In this case the image of the mirror C in A coincides with D. When the adjustment is altered there will be a path retardation between the

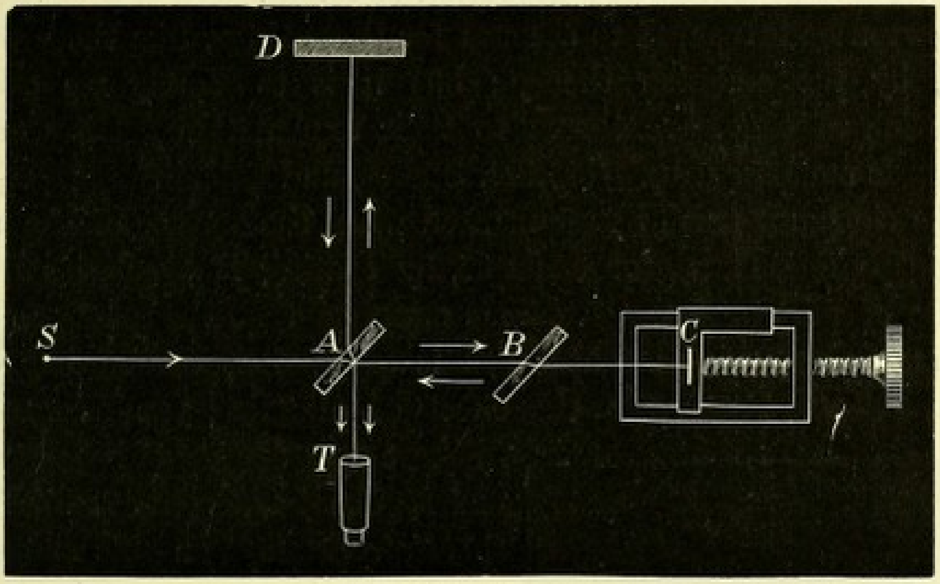



Fig. 105.

pencils equivalent to that of an air film of thickness AC - AD, and whose angle is equal to that between D and the image of C in A.

One of the mirrors is furnished with a micrometer screw by which it can be moved parallel to itself so as to introduce any difference of path desired.

**124. Newton's Diffusion Rings.** — The preceding cases of interference may be produced with tolerably thick plates and are known as the interference colours of thick plates. The coloured rings of thick plates are phenomena of a distinct kind. These were discovered and described by Newton.[1]

He allowed a pencil of sunlight to fall perpendicularly upon a glass mirror, ground concave on one side and convex on the other, to a sphere of nearly six feet radius, and silvered on the convex surface. Holding at the centre of the sphere a screen of white paper, with a small hole at its centre to allow the beam of light to pass and repass, he observed four or five rainbow-coloured rings on the paper, encircling the aperture through which the light poured. These were similar to the *transmitted* rings of thin plates, the squares of the radii of the bright rings being proportional to the even numbers, while those of the dark rings were in the ratios of the odd numbers.

When the mirror is inclined a little so as to throw the reflected image slightly to one side of the aperture the rings are formed as before, but their centre is at the middle point of the line joining the

[1] Newton, *Opticks,* book ii. part iv.

aperture and its image. This central spot changes its appearance in a remarkable manner as the image recedes from the aperture, being alternately bright and dark when homogeneous light is used, but with white light it assumes every gradation of colour.

Newton endeavoured to account for these rings, like those of thin plates, by the emission theory; but it is to Young we owe the application of the wave theory to their explanation. His work was afterwards completed by Herschel, Stokes, and Schäfli.[1]

The rings are much better marked when the surface of the mirror is slightly dimmed, as, for example, by coating it over with a weak mixture of milk and water. This suggests that the phenomena are due in some way to the light which is *scattered* at the surface of the mirror, for we know that the effect of dimming the surface is to increase the quantity of light scattered or irregularly reflected at it. We will therefore attempt on this supposition to explain the phenomena in the more general case when the light comes from a point L (Fig. 106) of the screen not far from the axis of the mirror, the plane of the screen passing through the centre O of the mirror and being perpendicular to its axis ON. Denote the distance LO by $u$, the radius OA of the first surface of the mirror by $a$ the radius ON of the second by $a + e$ where $e$ is the thickness of the mirror. Consider the illumination at any point Q (not in the plane of the paper) on the screen.

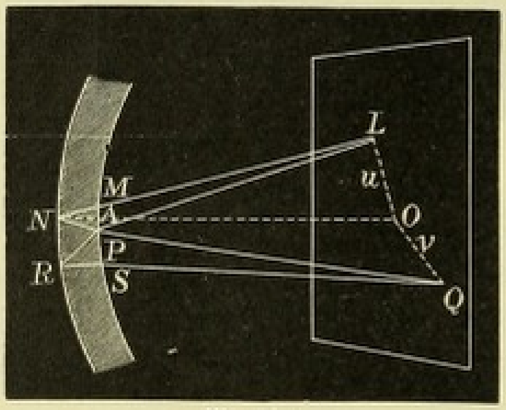


Fig. 106.

A ray of light LM, after incidence at M, will be refracted in part along MN, and after suffering reflection at N it will arrive at P, where a portion will be regularly refracted and some of it will be scattered. Let PQ be the scattered ray which reaches Q. Now another stream of light will also reach Q. For consider the ray incident at P, part of it will be scattered on entering the plate, and some one PR of these irregular rays will after regular reflection and refraction at R and S respectively reach Q along SQ. It will thus have traversed a path LPRSQ differing little from that traversed by LMNPQ. The two pencils will therefore be in a condition to produce interference effects.

[1] These rings are due to the interference of the light scattered or diffracted by the *same* particle of dust. It has been shown by Stokes that no regular interference is to be expected between portions of light diffracted by different particles of dust, for the diffusion is accompanied by a difference of path which varies from point to point (*Camb. Trans.* ix. p. 147, 1851). In this memoir there is a complete discussion of the whole case.

It may be observed that both the rays have suffered the scattering at the same point P of the surface, the first at emergence and the second at incidence. The two have therefore passed over similar paths and suffered similar transformations. They are consequently beams of the same nature, and may interfere.

It only remains now to calculate the difference of the paths traversed by the rays. We have

$$\mathrm{LM} = \sqrt{a^2+u^2} = a\left(1+\frac{u^2}{a^2}\right)^{\frac{1}{2}} = a + \frac{u^2}{2a} \qquad \text{(approx.)}$$

since $u$ is very small compared with $a$. Again, if $i$ be the angle of incidence (LMO) of the ray LM, its sine is approximately equal to its tangent or circular measure, and consequently

$$\sin i = \frac{\mathrm{OL}}{\mathrm{OM}} = \frac{u}{a},$$

and therefore

$$\sin r = \frac{\sin i}{\mu} = \frac{u}{\mu a}.$$

But $r = \mathrm{MNA}$, therefore

$$\mathrm{MN} = (e^2 + \mathrm{MA}^2)^{\frac{1}{2}} = e\left(1+\frac{\mathrm{MA}^2}{2e^2}\right) = e\left(1+\frac{u^2}{2\mu^2a^2}\right),$$

since

$$\frac{\mathrm{MA}}{e} = \sin \mathrm{MNA} = \sin r = \frac{u}{\mu a}.$$

Similarly

$$\mathrm{NP} = \mathrm{MN} \text{ and } \mathrm{PQ} = a + \frac{r^2}{2a}.$$

But the path MN + NP is traversed in glass, so that the equivalent path in air will be $\mu$ times as great; the whole equivalent air path is then

$$\mathrm{LM} + \mu(\mathrm{MN}+\mathrm{NP}) + \mathrm{PQ},$$

or

$$a + \frac{u^2}{2a} + 2\mu e\left(1+\frac{u^2}{2\mu^2a^2}\right) + a + \frac{v^2}{2a}.$$

From this the value of the path LPRSQ may be written down at once by observing that the first ray travels regularly till it emerges at P, where it is scattered. But if the second path be traversed in the reverse direction we see that, setting out from Q along QS, refraction and reflection take place regularly at S and R till emergence at P, where scattering takes place. Traversing the second path in the direction QSRPL is then the same as the first path in the direction LMNPQ with this difference, that the point Q is at a distance $v$ from O, whereas the distance of L from it is $u$. Consequently the second path may be written down from the first by interchanging $u$ and $v$. It is therefore

$$a+\frac{v^2}{2a}+2\mu e\left(1+\frac{v^2}{2\mu^2a^2}\right)+a+\frac{u^2}{2a}.$$

The difference of the path is consequently

$$\delta=\frac{e}{\mu a^2}(v^2-u^2).$$

When this difference is an even number of half wave lengths the two beams are in the same phase, and all points at the distance $v$ from O are bright. If the difference is an odd number of half waves this circle is dark.

There is therefore a series of alternately bright and dark rings encircling the point O. The radii of the rings are the values of $\rho$, which satisfy the equation

$$\frac{e}{\mu a^2}(\rho^2-u^2)=n\frac{\lambda}{2},$$

the bright rings corresponding to even values of $n$ and the dark rings to the odd values.

If $u=0$, we have then

$$\rho^2=\frac{\mu a^2}{e}\cdot n\frac{\lambda}{2},$$

a formula which embraces the laws laid down by Newton for the case in which the origin of light is at the centre of the mirror.

Thus the diameter of any ring varies inversely as the square root of the thickness of the mirror, while, as in the transmitted rings of thin plates, the diameters of the bright rings are proportional to the square roots of the even numbers, and the diameters of the dark rings to the square roots of the odd numbers. Finally, the squares of the radii vary directly as the wave length, so that with white light the circles are rainbow-coloured bands, changing from violet at the inner to red at the outer edge.

When $u=v$ the retardation is zero for all wave lengths, consequently the circle with centre O, and passing through the luminous point L, will be bright for all colours, and will therefore be white, and (opposite to L) at the other extremity of the diameter of this circle there will be an image of the point L formed by regular reflection at the surface of the mirror.[1] This white circle is surrounded by another

[1] With regard to these Newton observes: "The incident and reflected beams of light always fell upon the opposite parts of this white ring, illuminating its perimeter like two mock suns in the opposite parts of an iris" (*Opticks*, book ii. part iv. obs. 10).

system of coloured rings. The difference of the squares of the radii of any two consecutive rings is constant, and the area of the annulus between any consecutive pair is therefore the same for all pairs and is equal to

$$\pi(\rho_1^2 - \rho_2^2) = \frac{\pi\mu a^2}{e} \cdot \frac{\lambda}{2}.$$

The successive circles therefore approach each other more and more closely the farther we go out from the centre, so that at a short distance from the centre the appearances become so confused that all effect is obliterated. These rings may be also observed either with a plane glass mirror or with a convex mirror, but they are more faint. With a convex mirror it is necessary to use a convergent pencil of light in order to obtain a real image of the source.

They were obtained by the Duc de Chaulnes[1] with a concave metallic reflector, in front of which he placed a very thin plate of glass or mica dimmed with a mixture of milk and water. The metallic reflecting surface plays the part of the silvered back of Newton's mirror, while the plate of glass or mica acts as its first face, and the air space between the plate and the reflector corresponds to the glass of the mirror.

The rings may also be well observed without a screen, as suggested by Stokes.[2] All that is required is to place a small flame at the centre of curvature of the mirror, so that it coincides with its image. The rings are then seen surrounding the flame.

Dr. Whewell[3] observed a similar system of coloured bands formed when the image of a candle, held near the eye, is viewed by reflection in a plane glass mirror placed at a distance of some feet. This observation was communicated to M. Quetelet,[4] by whom it was published. In repeating the experiment together they found it essential that the mirror should not be perfectly bright, and to ensure the production of the bands it was sufficient to breathe gently on the surface of a cool mirror. Instead of moisture, which quickly evaporates, M. Quetelet recommended a tarnish of grease.

**125. The Colours of Mixed Plates.**—When the space between two glass plates is filled with a mixture of two substances in a finely divided state—such as water and air, or water and oil—light will in general traverse the different parts of the mixture in different times, and the interval of retardation will depend upon the difference of the velocities

[1] De Chaulnes, *Mém. de l'Acad. des Sc.* p. 136, 1755.
[2] Stokes, *Phil. Mag.* p. 419, 1851.
[3] Stokes, *Camb. Phil. Trans.* pp. 148, 149, 1851.
[4] Quetelet, *Corresp. Phys. et Math.* tom. v. p. 394, 1829.

of light in the two media, and upon the varying arrangement of the media between the two plates. Portions of the transmitted light will therefore be in a condition to interfere, and coloured rings are seen when a luminous object is viewed through the glasses.

The colours of mixed plates were first observed by Dr. Thomas Young and described in the *Philosophical Transactions* for 1802.[1] He produced them by interposing small globules of water, or butter, between two glass plates, or two object-glasses, pressed together so as to give the ordinary colours of thin plates. In this way little cavities of air were surrounded with water or butter, and on looking through the combination he saw fringes or coloured rings several times larger than those of thin plates which would have been produced had air alone been contained between the glasses. The rings were seen by the direct light of a candle, and began from a white centre like those produced by transmission through an air film. On the dark space next the edges of the plate he observed another system of fringes complementary to the first and beginning from a dark centre like those produced by reflection. This latter system was always brighter than the former. Brewster[2] in repeating the experiments "tried transparent soap and whipped cream, which gave tolerably good results; but I obtained the best effect by using the white of an egg beat up into froth. To obtain a proper film of this substance I place a small quantity between the two glasses, and having pressed it out into a film I separate the glasses, and by holding them near the fire I drive off a little of the superfluous moisture. The two glasses are again placed in contact, and pressed together so as to produce the coloured fringes or rings, they are then kept in their place either by screws or by wax, and may be preserved for any length of time."

If a dark object be behind the glasses and if the incident light be somewhat oblique, the rings change their character and resemble the ordinary reflected system of Newton. One of the portions of the interfering light in this case suffers reflection, and accordingly half a period difference of phase is introduced. These rings contract as the obliquity of the light is increased. The opposite occurs in the case of Newton's rings.

The colours of mixed plates were attributed by Young to interference, on the supposition that part of the transmitted light passes through one of the constituents of the mixture and part through the

[1] Also republished in 1807, *Elements of Natural Philosophy*, vol. i. pp. 470, 787; vol. ii. pp. 635, 680.

[2] Sir D. Brewster, "On the Colours of Mixed Plates," *Phil. Trans.* p. 73, 1838.

other, and the explanation from this point of view has been followed up by Verdet.[1] Thus if part of the incident light be refracted through the plate at an angle $r_1$, and another part at an angle $r_2$, then by Ex. Art. 71 the relative retardation between these parts is

$$\delta = e(\mu_1 \cos r_1 - \mu_2 \cos r_2),$$

and when this is an even or an odd number of half wave lengths there will be a corresponding increase or diminution of intensity. For normal incidence $\delta = e(\mu_1 - \mu_2)$, and the bright and dark rings will correspond to

$$2e(\mu_1 - \mu_2) = n\lambda,$$

according as $n$ is even or odd.

In the case of the coloured rings of thin plates seen by transmitted light the bright and dark rings correspond to thicknesses

$$2e' = n\tfrac{1}{2}\lambda,$$

according as $n$ is even or odd.

Hence the thicknesses $e$ and $e'$, at which rings of the same order ($n$) occur in the two experiments, are in the ratio

$$\frac{e'}{e} = \tfrac{1}{2}(\mu_1 - \mu_2).$$

## *Examples*

1. If the constituents of the mixed plate be water and air, we have approximately $\mu_1 = \frac{4}{3}$ and $\mu_2 = 1$, therefore

$$\frac{e'}{e} = \frac{1}{6}.$$

But the squares of the diameters of the rings are proportional to the thicknesses, consequently the diameters of corresponding rings of the two systems are in the ratio $1 : \sqrt{6}$.

2. The expression $\mu_1 \cos r_1 - \mu_2 \cos r_2$ increases with the angle of incidence, and the rings consequently contract in diameter. For

$$d\delta = e(\mu_2 \sin r_2 dr_2 - \mu_1 \sin r_1 dr_1) = e \sin i(dr_2 - dr_1),$$

but $dr_2$ is greater than $dr_1$, therefore $d\delta$ is positive and $\delta$ increases with $i$.

## *Newton's Observations on the Diffusion Rings of Thick Plates*

*Opticks*, book ii. part iv.: "There is no glass or speculum how wellsoever polished, but, besides the light which it refracts or reflects regularly, scatters every way irregularly a faint light, by means of which the polish'd surface, when illuminated

[1] Verdet, *Optique Physique*, tom. i. p. 155.

in a dark room by a beam of the sun's light, may be easily seen in all positions of the eye. There are certain phenomena of this scattered light which, when I first observed them, seem'd very strange and surprising to me. My observations were as follows.

"*Obs.* 1.—The sun shining into my darken'd chamber through a hole one-third of an inch wide, I let the intromitted beam of light fall perpendicularly upon a glass speculum ground concave on one side and convex on the other, to a sphere of five feet and eleven inches radius, and quick-silvered over on the convex side. And holding a white opake chart, or a quire of paper at the centre of the spheres to which the speculum was ground—that is, at a distance of five feet and eleven inches from the speculum—in such a manner, that the beam of light might pass through a little hole in the middle of the chart to the speculum, and thence be reflected back to the same hole: I observed upon the chart four or five concentric irises or rings of colours, like rainbows, encompassing the hole much after the manner that those which (in the fourth and following observations of the first part of this third book) appear'd between the object-glasses, encompassed the black spot, but yet larger and fainter than those. . . . When the sun shone very clear there appeared faint lineaments of a sixth and seventh. If the distance of the chart from the speculum was much greater or much less than that of six feet, the rings became dilute and vanished. And if the distance of the speculum from the window was much greater than that of six feet, the reflected beam of light would be so broad at the distance of six feet from the speculum where the rings appeared, as to obscure one or two of the innermost rings. And therefore I usually placed the speculum at about six feet from the window; so that its focus might fall in with the centre of its concavity at the rings upon the chart."

"*Obs.* 3.—Measuring the diameters of these rings as accurately as I could, I found . . . the squares of the diameters of the (bright) rings in the progression, 0, 1, 2, 3, 4, etc. I measured also the diameters of the dark circles between these luminous ones, and found their squares in the progression of the numbers $\frac{1}{2}$, $1\frac{1}{2}$, $2\frac{1}{2}$, $3\frac{1}{2}$, etc." (that is 1, 3, 5, 7, etc.).

"*Obs.* 5.—. . . If the speculum was illuminated with red (light), the rings were totally red with dark intervals, if with blue they were totally blue, and so of the other colours. . . . In the red they were largest, and the indigo and violet least, and in the intermediate colours, yellow, green, and blue, they were of several idtermediate bignesses."

"*Obs.* 7.—By analogy . . . it seemed to me that these colours were produced by this thick plate of glass much after the manner that those were produced by very thin plates. For upon trial I found that if the quicksilver were rubbed off the backside of the speculum, the glass alone would cause the same rings of colours, but much more faint than before; and therefore the phenomenon depends not upon the quicksilver, unless so far as the quicksilver by increasing the reflexion of the backside of the glass increases the light of the rings of colour. I found also that a speculum of metal without glass . . . produced none of these rings, and thence I understood that these rings arise not from one specular surface alone, but depend upon the two surfaces of the plate of glass whereof the speculum was made, and upon the thickness of the glass between them."

Experimenting with specula of different thicknesses, Newton found that the squares of the radii of the rings varied inversely as the thickness of the glass (see Observation 9).

"*Obs.* 12.—When the colours of the prism were cast successively on the speculum, that ring which in the two last observations was white, was of the same bigness in all the colours, but the rings without it were greater in the green than in the blue, and still greater in the yellow and greatest in the red. And, on the contrary, the rings within that white circle were less in the green than in the blue and less in the yellow and least in the red."

# CHAPTER IX

## DIFFRACTION

### Section I.—The Elementary Theory

**126. Diffraction—Introductory.**—By far the greatest difficulty encountered at the outset of the wave theory was the explanation of the rectilinear propagation of light. This difficulty confronted the early founders of the theory, and those who feared to encounter it became partisans of the emission theory from which the principle flowed as a natural consequence. A closer examination of the facts, however, shows that light suffers some deviation from the rectilinear course in passing by the edge of an opaque obstacle; that it does bend round corners as required by the wave theory, but, as the wave length is excessively small, the intensity falls off rapidly within the geometrical shadow, and the amount of bending observable is so slight that careful examination is required to detect it. When light passes through a very small aperture, the dimensions of which are comparable with the wave length, it is not propagated through the aperture as a definite ray or pencil, but diverges in all directions just as sound does when passing through an aperture of a few feet in diameter. Bending.

On the other hand, well-marked sound shadows are formed by huge obstacles, such as mountains or large buildings, and the existence of these may be often noticed by the most casual observers.

The phenomena which occur when light passes through a very narrow aperture or close to the edge of an opaque obstacle, and which arise from the light deviating from the rectilinear path, are classified and studied under the title of *Diffraction*. These appearances are depicted at the boundaries of the geometrical shadow, and while all other theories have failed to account for them, the wave theory explains and even predicts the phenomena truly.

The first observations on diffraction were made by Grimaldi,[1] about

[1] *Physico-mathesis de lumine, coloribus et iride*, Bononiæ, 1665.

Grimaldi. the middle of the seventeenth century. Having admitted a cone of light into a darkened chamber through a very small aperture, he found that when a small opaque obstacle was placed in the cone its shadow on a screen was much larger than its geometric projection, so that the light suffered some deviation from the rectilinear course in passing the edge of the obstacle. On observing the shadows attentively he found that they were bordered by three iris-coloured fringes, running parallel to the edge of the shadow, and decreasing in width and intensity as their distance from it increased. Similar fringes may also be observed, under favourable conditions, within the shadows of narrow obstacles such as a fine wire or hair.

Newton. The phenomena of diffraction were subsequently examined by Hooke[1] and Newton.[2] The experiment of Grimaldi was varied by Newton, who transmitted light through a very narrow aperture between two knife edges and observed the image it cast upon a screen. The image was bordered by three iris-coloured bands, in which the colours succeeded each other as in the rings of thin plates—violet nearest the shadow and red farthest removed from it. He observed the same appearances in the exterior of the shadows of many obstacles, but he does not mention the brilliant fringes which occur in the interior of the shadows of narrow obstacles, although Grimaldi had observed the crested fringes at the angles of shadows.

Young. The first application of the wave theory to the explanation of diffraction phenomena was made by Dr. Young,[3] who attributed the fringes to the interference of the direct light which passes very close to the edge with the light reflected by the edge at grazing incidence. That the effects are not produced by the interference of two pencils of light, such as Young imagined, is proved by the fact that they do not depend on the degree of polish or sharpness of the edge. Fresnel observed that whether light passed over the back or edge of a razor the fringes produced were the same, and this he confirmed by exact experiments.[4] If the fringes depended in any way on the light reflected at the edge of the obstacle, they should vary in intensity or position when the degree of polish or the material of the edge is altered.

To ascertain whether the form of the edge had any effect on

[1] Hooke, *Posthumous Works*, p. 190, London, 1705.

[2] Newton, *Opticks*, book iii.

[3] Young, "On the Theory of Light and Colours," *Phil. Trans.* p. 12, 1802; *Lectures on Natural Philosophy*, pp. 342, 365, 1807.

[4] Fresnel, *Œuvres complètes*, tom. i. pp. 148, 280.

the fringes, Fresnel took two plates of steel, the edge of each being rounded off in one half of its length, but sharp in the remaining half. He placed the rounded portion of each opposite the sharp part of the other. If the position of the fringes depended on the sharpness of the edge the effect would here be doubled, and the fringes would appear broken in the midst. On the contrary, they were found perfectly straight throughout their entire length. Young's explanation is therefore incorrect, and it is to Fresnel that we owe the true solution of the problem, and the deduction of general expressions for the effect of a wave at any point. Fresnel.

According to Fresnel the phenomena of diffraction are to be attributed to the mutual interference of the secondary wavelets which diverge from the wave front. Each element of the primary wave, when it reaches the obstacle, is considered as the centre of a diverging wavelet, and the resultant of all the secondary waves he expressed by means of two integrals taken within limits determined by the particular nature of the problem under consideration. The problem of diffraction was thus solved by the principle of Huygens combined with the principle of interference.[1]

Diffraction phenomena are, therefore, due to the mutual interference of the disturbances propagated from the various elements of a single wave, just as the interference phenomena described in the foregoing chapters are due to the mutual interference of two trains of waves. The bright focus of a lens exists because the disturbances propagated there by the wave passing through the lens agree in phase and produce an intense effect. At other points destructive interference of the wavelets occurs, and there is no illumination.

In the hands of Fresnel[2] the theory was developed to such perfection that little room was left for addition, and by the exact agreement of the results of careful observation with the anticipations of analysis, the evidence furnished in its favour is so overwhelming that, to those who impartially examine it, little doubt is left as to its truth.

In this section we shall confine ourselves to a general explanation of the phenomena by elementary methods, so that we may become acquainted with the general facts of some important cases before entering upon the more intricate investigations.

[1] Marian (*Mém. de l'Académie des Sciences*, p. 53, 1738) and Dutour (*Mém. de l'Académie des Sciences*, tom. v. p. 365, 1768) conceived that the fringes depended on the refraction of the light by a thin layer of condensed air on the edge of the obstacle, but that this is not the case is proved by the fact that diffraction takes place in vacuum exactly as in air.

[2] [It appears from a note of Prof. Preston's that his opinion of Fresnel's work on diffraction was considerably modified after the publication of the second edition.]

**127. A Straight Edge.**—The first case of diffraction which we shall consider is that presented when light, diverging from a luminous point (such as the image of the sun produced by a lens of short focal length), passes by the straight edge of an opaque screen. Let O (Fig. 107) be a luminous point emitting spherical waves, AB an opaque obstacle perpendicular to the plane of the paper, then if the light be propagated accurately in right lines from O there should be uniform illumination above the line OAM, and complete darkness below it on a screen PQ placed to receive the light after it passes over the obstacle AB. It is observed, however, that the illumination does not become zero immediately below M, but that it fades away *continuously* and rapidly, and there is complete darkness at a small distance below M. Immediately above M, on the other hand, the illumination is not uniform, but passes through a series of maxima and minima, giving rise to a series of brilliant fringes parallel to the edge of the obstacle AB. These fringes become less distinctly marked the farther we recede from M, the boundary of the geometrical shadow, till at length they are wholly obliterated and merge into uniform illumination at a short distance from M.

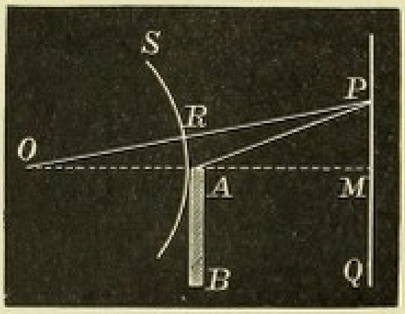


Fig. 107.

The shadow is thus not distinctly marked by the line OAM, as the geometrical theory of optics would indicate, but the light fades away gradually on one side, and passes through many alternate successions of brightness and darkness, forming fringes, on the other. The appearance of these fringes is independent of the distances of the two screens from the luminous point, the scale merely varying according to circumstances, and from their constant character they may be readily recognised in experiments in which they occur associated with other fringes.

Let us now consider the illumination at any point P on the screen outside the geometrical shadow. We have already found (chap. iii.) that the effect of the wave SA at P is confined to a limited number of half-period elements around the pole R. If therefore P is so far removed from M that none of the effective elements of the wave at P are intercepted by the screen AB, then the illumination at P will be affected in no way by the obstacle AB. But if P be so near M that the arc RA contains only a portion of the effective elements of the lower half of the wave, the other part being intercepted by the obstacle, we may roughly consider the illumination at P as consisting of two portions, one the entire half wave RS and the other the part

RA consisting of a few half-period elements. Now if RA contains an even number of these elements they will mutually interfere in pairs and should have little effect at P. Whereas if RA contains an odd number the effect should be much greater at P. We are led to infer then that the minimum at P is a maximum or a minimum according as the arc RA contains an odd or an even number of half-period elements. That is, if

Fringes outside.

$$AP - RP = (2n+1)\frac{\lambda}{2} \qquad \text{(maximum brightness)}$$

and if

$$AP - RP = 2n\frac{\lambda}{2} \qquad \text{(minimum brightness)}$$

the difference of the distances AP and RP remaining constant, the point P will move along a hyperbola having A and O for foci, for

$$OP - AP = OR - (AP - RP),$$

and OR is constant, therefore the difference of the distances of P from the fixed points O and A is constant. P therefore describes a hyberbola, but its curvature is so small that it almost coincides with its asymptotes.

Denoting OA by $a$ and AM by $b$ we can easily calculate the distance $x = PM$ of any bright or dark band from M; for[1]

$$OP = a + b + \frac{x^2}{2(a+b)}, \quad \text{and } AP = b + \frac{x^2}{2b}.$$

Therefore

$$AP - RP = \frac{x^2}{2}\left(\frac{1}{b} - \frac{1}{a+b}\right),$$

or

$$\frac{x^2}{2}\frac{a}{b(a+b)} = n\frac{\lambda}{2} \qquad \text{(for a bright or dark band).}$$

Hence for maximum brightness

$$x = \sqrt{\frac{b(a+b)}{a}(2n+1)\lambda},$$

and for minimum brightness

$$x = \sqrt{\frac{b(a+b)}{a} \cdot 2n\lambda}.$$

---

[1] Or thus: the angle PAM is very small and equal to twice the angle ARD, where D is the point in which a circle, centre P, radius PR, cuts PA, therefore

$$\frac{x}{b} = 2\frac{AD}{AR} = 2\frac{\delta}{a \,.\, \angle AOR} = 2\frac{\delta}{a} \div \frac{x}{a+b}, \quad \text{or } x^2 = 2\frac{b(a+b)}{a} \,.\, \delta.$$

When the screen PM is moved nearer to or farther from the obstacle AB, the distance $a$ remains constant while $b$ varies; the corresponding values of $x$ are the ordinates of the hyperbola mentioned above.

To determine the illumination at any point P inside the geometrical shadow, we must observe that the part of the wave which propagates light to P is only a fraction of half a wave. The point R (Fig. 108) is now below the edge of the obstacle, so that some of the powerful elements are intercepted by the screen. The light from the point A is, however (of the effective part of the wave), that which requires least time to reach P. If therefore we divide the part AS of the wave into half-period elements with respect to P beginning at A, the first element $AM_1$ will be the most powerful, and the others will be smaller and become rapidly equal to one another, so as to destroy each other's effect at P. The resultant effect at P is therefore confined to a few half periods near the edge A of the obstacle, and these will give a resultant illumination at P. However, as P sinks farther into the shadow, the obliquity of the line PA to the wave front will increase, and the consecutive half-period elements $AM_1$, $M_2$, $M_3$, etc. with respect to P will gradually become smaller and more nearly equal in effect, till finally when P is a small distance below M, the whole effective portion of the wave is cut off and the resultant at P is zero.

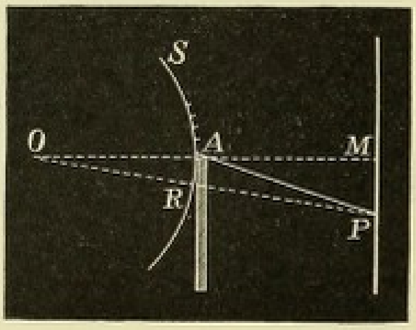


Fig. 108.

No fringes inside.

Consequently we conclude that the light falls off continuously but rapidly within the geometrical shadow, and alternations of brightness and darkness do not occur.

It is not difficult to see that diffraction fringes can be exhibited only when the angular diameter of the source of light is small. For if the luminous origin subtends any considerable angle at the eye, each point of it will give rise to a corresponding set of fringes, and the multitude of sets will be so superposed and intermixed as to obliterate all visible effect.

In practice a strip of light from a narrow slit is used, and the fringes are viewed through an eyepiece mounted on an optical bench.

**128. Narrow Wire.**—Let us now consider the case of a very narrow opaque obstacle, such as a hair or a fine wire. Let the opaque screen AB of the preceding article be limited on the lower side, so that it becomes a narrow strip intercepting the light from O. The shadow on the screen MN (Fig. 109) will be bounded externally on each side by a system of fringes similar to those just described, and accordingly attributable to the same cause. The system on either side is produced by the light which passed that side of the obstacle acting independently of that which passed the other side. The upper system is due to the diffraction of the light from O over the upper side A of the obstacle, and the lower system is produced by the diffraction of the light at the lower side B.

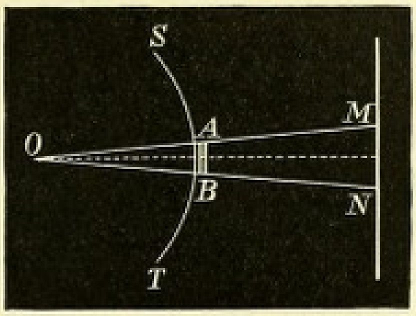


Fig. 109.

In addition to these two systems of fringes, however, there is another set of brilliant bands situated inside the geometrical shadow (MN) of the obstacle (if it be sufficiently narrow). This system is finer than the others, and, unlike them, of equal width throughout. It remains therefore to account for this internal system by the theory.[1]

In the preceding article we have seen that the effect of the portion AS of the wave at any point P inside the geometrical shadow is due entirely to a few half-period elements at A, the others mutually destroying each other. The portion AS of the wave may therefore be replaced by a small luminous source near A, so far as its effect in illuminating P is concerned. Similarly the lower portion BT of the wave, in illuminating P, is equivalent to a small luminous source near B. Thus any point P within the shadow is illuminated by both sources at the same time; these sources will, therefore, like two small near apertures, produce the phenomena of interference inside the shadow, and the fringes which occur there are accordingly accounted for.

If the obstacle AB is not very narrow, then there will be no internal fringes, but a gradual fading away of the light at each side of its geometrical shadow and the usual system of external diffraction fringes on each side. When the obstacle is narrow, however, the illumination inside M overlaps that inside N, and interferes with it. Or we might put it thus: with the straight edge any point inside the shadow is illuminated by a small source at A. If the obstacle be narrow this point is also illuminated by a small source at B, and if the distance AB between these sources is small enough, they will interfere and produce fringes

[1] Diffraction fringes may be exhibited on a minute scale by candle light, with no other apparatus than a small lens having a fine wire stretched across in contact with its surface. Holding the other surface next the eye, if we look through the lens at the flame of a candle at some distance, or, what is better still, at its light admitted through a narrow slit, the wire being parallel to the slit, the dark image of the wire will be seen edged by the external fringes, and the shadow marked by the internal fringes in a remarkably beautiful and distinct manner (see B. Powell, *Phil. Mag. and Ann.* January 1832).

in the interior of the geometrical shadow. It is clear that these fringes are given by the equations

$$AP - BP = n\frac{\lambda}{2}, \quad \text{and } x = \frac{a}{c}n\frac{\lambda}{2},$$

where $n$ is even for the bright bands and odd for the dark ones.

That the internal fringes are due to the interference of the two portions of light which pass over the edges of the narrow obstacle was proved conclusively by Dr. Young. He showed that when the light from one side was intercepted by an opaque screen either before or after it reached the obstacle, the whole system of internal fringes disappeared, and the ordinary external diffraction fringes alone remained on that side over which the light was allowed to pass. It is clear therefore that the internal fringes are due to the joint working of the light which passes both sides of the obstacle, whilst the external fringes on the upper and lower sides of the shadow are due to the independent action of the portion of the light which passes on these sides respectively.

**129. Narrow Rectangular Aperture.**—Let us now turn to the quasi-complementary case,—that in which the light from a source O is admitted to a screen through a very narrow slit or rectangular aperture.[1] First consider the illumination at any point P (Fig. 110) of the screen outside the boundary of the geometrical image of the aperture. As before, we divide AB into a series of half-period elements with respect to P, beginning at A, as the light reaches P from this point first. Then the point P will be the centre of a bright or dark band according as the difference BP – AP is an odd or an even number of half-wave lengths. If the difference is equal to an even number of half waves there will be an even number of half-period elements in AB, which will mutually interfere at P, and the effect will be less than

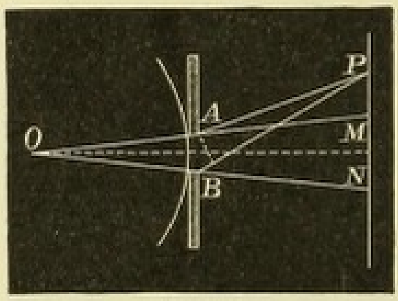


Fig. 110.

[1] The fringes of an aperture may be very distinctly seen by merely placing a narrow slit near the flame of a candle, and viewing it through another slit held close to the eye and parallel to it at the distance of a few feet.

Diffraction spectra, images, or patterns may, however, be observed without the aid of any subsidiary apparatus, by partially closing one's eyes so as to view through the eyelashes a candle flame or any ordinary source of light. The lashes in this case play the part of a diffraction grating, and spectral images are produced on each side of the flame. Similar phenomena may be observed by viewing any ordinary source of light, such as a lamp flame or a gas jet, through a pocket-handkerchief. These effects are best marked when the edge of the flame is presented to the observer so that it may have its smallest angular subtense.

A star which is probably a vast body gives diffraction rings as a point source because its angular subtense is very small.

if there is an odd number of half-period elements in AB. There will thus be two systems of fringes, one on each side of the geometric image of the aperture, the bright and dark bands of which correspond to odd and even values of $n$ respectively. This is just the reverse of what takes place in the fringes formed in the interior of the shadow of an opaque obstacle. The position of these fringes is given as usual by the equations

$$BP - AP = n\frac{\lambda}{2}, \quad \text{and } x = \frac{a}{c} n \frac{\lambda}{2},$$

where $n$ is even for the dark and odd for the bright bands.

Now if the screen is so remote from the aperture that AM − BM is less than half a wave, then the first band will lie outside the edge of the image, and the systems of fringes already mentioned will represent the complete phenomena. But if the screen is so near the aperture that the difference of the distances of M (or N) from A and B is a number of half waves, fringes are visible within the projection of the aperture also. The illumination at any point Q of this image is due to the two portions into which OQ divides the wave AB. These portions are sensibly different in magnitude as well as obliquity, and their joint effect at Q requires a more complete investigation (Art. 150).

**130. Circular Aperture.**—Among the most striking of the phenomena of diffraction are those produced when light, diverging from a luminous origin, passes through a small circular aperture such as a pinhole in a sheet of lead. When the aperture is viewed through a lens it appears as a brilliant spot surrounded by a series of vivid rings, and as the distance between the aperture and the eye is altered these rings vary in the most beautiful manner. The central white spot contracts to a point and vanishes as the eye approaches the aperture while the rings close in upon it in succession, and the centre passes in succession through a series of most beautiful hues similar to those presented in the colours of thin plates.

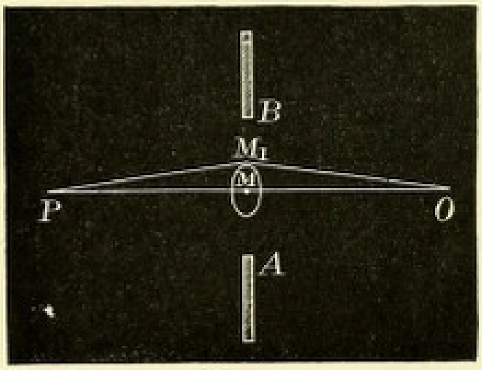


Fig. 111.

The points of maximum and minimum intensity on the central line are easily determined. Thus let O be the luminous origin (Fig. 111), AB a section of the aperture, and OMP a line through its centre M. The illumination at P is found by dividing the wave into half-period elements as it diverges through AB. Thus let $M_1$, $M_2$, $M_3$, etc., be a series of points on the wave front marking the half-period elements. Then, as shown in Art. 53, the consecutive zones approximately destroy each other, and if the aperture transmits an even number of them, the illumination at P will be very feeble, whereas if it transmits an odd number the illumination will be largely increased. Hence as P travels along the axis of the aperture the intensity of the illumination at it passes through a succession of maxima and minima. The distance of any point of maximum or minimum intensity from the aperture is easily calculated. For if we denote OM and PM by $a$ and $b$ respectively it follows at once from the expression of Art. 53 that the area of each half-period element is very approximately equal to

$$\frac{\pi ab\lambda}{a+b},$$

since the radius of the aperture is small compared with either $a$ or $b$. Denoting this radius by $r$, the area of the aperture will be $\pi r^2$, and if this contains $n$ half-period elements we have approximately

$$\pi r^2 = \frac{\pi abn\lambda}{a+b}.$$

The positions of the points of maximum and minimum intensity along the axis are consequently given approximately by the equation [1]

$$b = \frac{ar^2}{na\lambda - r^2},$$

where $r$ is the radius of the aperture. The dark points correspond to the even values of $n$, and the brightest points to the odd values.

Since the distance of any bright point from M depends on $\lambda$, it follows that when white light is used points of maximum brightness for the different colours will be situated at different distances from M, the red being nearest M and the violet farthest away; there will not,

Dispersion of maxima.

[1] This result may be easily deduced directly by expressing the path retardation of the ray PAO, which passes the edge of the aperture, relatively to the ray PMO which passes through the centre. Thus since $r$ is small we have

$$\text{OA} = a + \frac{r^2}{2a}, \quad \text{and PA} = b + \frac{r^2}{2b}.$$

Consequently the retardation is

$$\delta = (\text{OA} + \text{PA}) - (a+b) = \frac{r^2}{2}\left(\frac{1}{a} + \frac{1}{b}\right).$$

That is

$$r^2 = \frac{2ab\delta}{a+b}$$

where $\delta = \frac{1}{2}n\lambda$ when the aperture contains a whole number of half-period elements. This expression also gives $\pi ab\lambda/(a+b)$ as the approximate area of each half-period element.

therefore, be any points of complete darkness, but the centre of the image will pass through a succession of richly coloured hues, following each other nearly in the order of Newton's scale.

The dark points correspond to the even values of $n$, but a closer calculation shows that in this case, as in the case of a rectangular aperture, the points of maximum brightness are not exactly half-way between the points of darkness. Both these questions will therefore be resumed in the next chapter, and be more fully dealt with.

**131. Zone Plates.**—We have seen that the consecutive annuli into which the circular aperture AB is divided approximately destroy each other in pairs at P. If then the alternate annuli, *e.g.* the 2nd, 4th, 6th, etc., be covered with some opaque substance, the others would be left free to have their full effect, and we should expect P to be brightly illuminated. The theory here is in complete accordance with the results of observation. Such plates, with alternately opaque and transparent annuli, may be obtained by photography, and it is found that, although there is an important difference, they resemble a lens in bringing light from any orgin O to the same point P as focus. The difference is that the light which passes through the second transparent annulus arrives at P a complete period later than the light from the first. Similarly the path of the light from the third is a wave length greater than the path of the light from the second, and so on. Whereas in the case of a lens the characteristic is that all the light which reaches the focus from O arrives there in the same time. The phases of all the waves which reach the focus are the same, and the times occupied by them in travelling there from O are also the same for all. A zone plate has therefore the property of a condensing lens.[1]

To construct a zone plate it is only necessary to describe on a slip of white paper a series of concentric circles of radii proportional to the square roots of the natural numbers, *i.e.* 1, $\sqrt{2}$, $\sqrt{3}$, $\sqrt{4}$, etc. The areas of these circles are directly as the natural numbers, and the area included between any pair of consecutive circles is constant, hence if the alternate annuli be blackened over, the others remaining white, we will have a sketch of a zone plate, but on much too large a scale. A miniature photograph of it may now be made on a thin plate of glass, and it is found that when this plate is interposed in the path of a beam of light it produces the effect described.[2] The light is

[1] A fourfold effect would be obtained if it were possible to provide that the light stopped by the alternate zones were replaced by a phase-reversal without loss of amplitude.

[2] See Glazebrook's *Physical Optics*, p. 182.

brought to a focus at a point P such that the rings on the plate are half-period elements with respect to it. The focus of the red light is of course, as before, nearer the plate than the focus of the violet light. The reverse holds in the case of a lens, for the violet, being most refrangible, comes to a focus nearest the lens.

**132. Opaque Circular Disc.**—In applying the theory to the case of diffraction by an opaque circular disc, Poisson was led to the startling conclusion that the illumination at the centre of the shadow should be the same as when the disc is removed, and Arago showed that this was verified by experiment.[1] Without entering into a complete investigation we can see that the illumination along the axis of the disc should be uniform and approximately the same as when the disc is removed. For, take any point on the axis of the disc, and divide the wave into half-period elements with respect to it, beginning at the edge of the disc. The first half-period element, which immediately surrounds the edge of the disc, is the most powerful, and plays the part ordinarily taken by the centre zone of the wave. As in chap. iii., it follows that the resultant effect is approximately equal to half that of the first existing zone, and when the obliquity is very small this will be very nearly the same as $\frac{1}{2}m$ (Art. 52), which represents the effect of the whole wave. The illumination, however, must fall off gradually as the point under consideration approaches the disc, for when the point is near the disc the obliquity of the first zone passing the edge is greater than when the point is farther away. In other words, the effective portion of the wave is to some extent cut off when the point is near the disc.

*Uniform intensity along axis.*

This result is easily understood by remembering the bending, or diffraction, of the light which takes place into the geometrical shadow. When the obstacle is a small circular disc this diffracted light overlaps from all sides at each point on the axis. At any one of these points all the diffracted light is in the same phase, and there is consequently no destructive interference at any point on the axis. At a point not on the axis the components of the diffracted light differ in phase, and there is interference, so that, when white light is used, what is observed is a system of rings surrounding a white centre, the intensity

[1] This experiment is difficult to perform satisfactorily, since even when the disc is cut with the utmost care, each of the minute inequalities in its edge is magnified and accompanied by fringes which mix and cross so as to totally confuse the whole appearance. "I have succeeded by taking up a small quantity of thick ink on the point of a pen, and dropping it on a clear plate of glass, by which means a sufficiently even circular edge is produced, the disk being about $\frac{1}{10}$ of an inch in diameter" (Rev. B. Powell, *Phil. Mag. and Ann.* January 1832). The difficulty may often arise from the faulty nature of the glass to which the disc is attached.

at the centre being practically the same as if the disc were removed.

**133. Babinet's Principle.**—When light is transmitted through a very small aperture we have seen that there will be illumination at points considerably outside its geometrical image. But when the aperture is of sensible magnitude, an image of it is depicted on the screen, and at the borders of this image the illumination falls off gradually. Now at any point outside the image, or at any point of the pattern where the illumination is zero, the effect produced by any part or parts of the aperture must be exactly equal and opposite to the effect of the remainder of the aperture. Thus if $S_1$ be the area of any portion or any number of portions, isolated or continuous, of the aperture, $S_2$ the area of the remainder, and S the whole area, we have

$$S = S_1 + S_2,$$

and if the vibration produced by $S_1$ at P, a point *outside* the image, be $y_1 = a \sin\phi$, then the vibration produced by $S_2$ will be $y_2 = -a \sin\phi$, for the resultant is zero. Hence it follows that if any part or parts $S_2$ of the aperture be supposed opaque, the remainder $S_1$ being left transparent, so that diffraction may occur when light is transmitted through the transparent parts, the vibration and illumination at P will be determined by the equations

$$y_1 = a \sin\phi, \quad \text{and } I_1 = a^2,$$

while if the foregoing opaque portions $S_2$ be supposed transparent, and the transparent portions $S_1$ opaque, the illumination at P when the light is transmitted through $S_2$ will be determined by

$$y_2 = -a \sin\phi, \quad \text{and } I_2 = a^2,$$

that is, $I_1 = I_2 = a^2$, or the illumination is *unaltered* when we suppose the transparent parts of the aperture to become opaque and the opaque parts transparent.

This principle is due to Babinet,[1] and applies to any point at which the illumination due to the whole aperture is zero.

If the point P be illuminated, then the whole aperture S transmits a vibration represented by the equation

$$y = A \sin(\omega t + \alpha),$$

while any selected portion $S_1$ of it produces the vibration

$$y_1 = a_1 \sin(\omega t + \alpha_1),$$

[1] Babinet, *Comptes Rendus*, tom. iv. p. 638, 1837.

and the remainder of the aperture produces

$$y_2 = a_2 \sin(\omega t + \alpha_2).$$

The vibrations due to the portions $S_1$ and $S_2$, in general, will differ in phase and amplitude, but they will be connected with the vibration due to the complete aperture S by the equation (Art. 43)

$$A^2 = a_1^2 + a_2^2 + 2a_1a_2 \cos(\alpha_1 - \alpha_2),$$

or denoting the corresponding illuminations by I, $I_1$, and $I_2$, we have

$$I = I_1 + I_2 + 2\sqrt{I_1 I_2} \cos(\alpha_1 - \alpha_2).$$

If $\alpha_1 - \alpha_2 = \frac{1}{2}\pi$, then $I = I_1 + I_2$. It should be remarked therefore that we have the equation $I_1 + I_2 = I$ only in the special case when the vibrations from $S_1$ and $S_2$ differ in phase by a quarter period. Attention is directed to this point, for it is not unusual to find it assumed that $I = I_1 + I_2$ universally, or that the illumination in one case is exactly complementary to that in the other.[1]

When we know the vibration $y$ produced by the whole aperture, and the vibration $y_1$ produced by any part, we can calculate by the above equations the vibration $y_2$, and illumination $I_2$, produced by the remainder. Thus if the aperture S be supposed very large, so that practically the whole wave may reach any point P on the screen, then if $y$ be the vibration excited at P by the complete wave, $y_1$ that when any area $S_1$ of the wave is stopped by an opaque obstacle, $y_2$ that of the remainder, viz. the vibration due to that part of the wave which passes the obstacle, then we have

$$y = y_1 + y_2.$$

The vibration $y$ will be practically uniform all over the screen, so that we have

$$y_1 + y_2 = \text{constant},$$

but $I_1 + I_2$ is not constant, for we have seen that $I_1 + I_2$ is not equal to I except in a very special case. If $y_1$ be the vibration at any point when light passes through a narrow rectangular aperture, and $y_2$ the vibration at the same point when the aperture is replaced by a wire of the same width, then $y_1 + y_2 = y$, which is constant over the screen, but we cannot say that the illuminations at any point are complementary

[1] [These results are well illustrated by Fig. 10 on p. 46 if we suppose the circle C deleted and the angle AOB variable. OP, OQ, OR represent $y_1$, $y_2$, and $y = y_1 + y_2$ respectively, and OA, OB, OC are proportional to $\sqrt{I_1}$, $\sqrt{I_2}$, and $\sqrt{I}$. If $I = I_1 + I_2$ the angle AOB (the phase difference) must be right. If $\sqrt{I_1}$ and $\sqrt{I}$ are known as well as the difference BOC of the corresponding phases, $\sqrt{I_2}$ is found as being proportional to CA or OB.]

in the two cases. The same remark applies to the case of a circular aperture and an opaque disc of the same dimensions.[1]

Such screens are called *complementary screens*, but the term must not be understood to refer to any complementary relation between the illuminations; it merely signifies that the transparent portions of one screen are replaced by opaque parts in the other, and *vice versâ*. In the case of an opaque disc we have seen that the illumination is approximately uniform along the central line, while in the case of the circular aperture it passes through a series of maxima and minima. So again it is the outside of the image of a large obstacle that is bordered with diffraction fringes, whereas in the case of the corresponding aperture the fringes lie within the geometrical image and the patterns are not complementary. Quasi-complementary cases.

**134. Coronas—Young's Eriometer.**—We have already seen that when light, diverging from a luminous point, passes by the edges of an opaque obstacle, systems of coloured fringes are formed parallel to the edges of the shadow. In the case of a circular disc, or a circular aperture, the fringes form a system of concentric circular rings. Instead of a single aperture if we have a large number of irregularly distributed small equal apertures (or of a large number of equal circular discs) in an opaque screen it may be shown (see Art. 158) that the diffraction pattern is the same as that produced by a single aperture multiplied in intensity by the number of apertures.

Instead of opaque discs we might equally have small regular globules of condensed vapour, as in a cloud, and it is to the diffraction by these globules that the coloured rings seen around the sun and moon, when observed through a thin cloud, are due. These rings are observed close to the surface of the sun and moon in hazy weather, and must not be confused with the larger rings or halos formed at some distance away. The halos are often seen in northern latitudes, and are due to ice crystals floating in the atmosphere, the angular radius of the first being from 22° to 23°. Coronas. Halos.

Newton observed coloured rings around both the sun and moon, and he was the first to attribute them to the action of water globules in the air. He describes them as follows (*Opticks*, book ii. part iv.):—

"For in June 1692 I saw by reflexion in a vessel of stagnating water three halos, crowns, or rings, or colours about the sun, like three little rainbows, concentrick to his body. The colours of the first or innermost crown were blue next the sun, red without, and white in the middle between the blue and red . . . these crowns

[1] If $I_1$ corresponds to a circular aperture and $I_2$ to a disc, then (Art. 132) $I_2 = I$, and $I_1 + 2\sqrt{I_1 I_2}\cos\delta = 0$, or $\cos\delta = -\frac{1}{2}\sqrt{I_1/I}$.

enclosed one another immediately, so that their colours proceeded in this continual order from the sun outward. . . . The like crowns appear sometimes about the moon, for in the beginning of the year 1664, February 19th, at night, I saw two such crowns about her. The diameter of the first or innermost was about three degrees, and that of the second about five degrees and a half. . . . At the same time there appeared a halo about 22° 35′ distant from the centre of the moon. It was elliptical, and its long diameter was perpendicular to the horizon, verging below farthest from the moon. I am told that the moon has sometimes three or more concentric crowns of colours encompassing her next about her body. The more equal the globules of water or ice are to one another, the more crowns of colours will appear, and the colours will be the more lively. The halo at the distance of 22°½ from the moon is of another sort. By its being oval and remoter from the moon below than above, I conclude that it was made by refraction in some sort of hail or snow floating in the air in a horizontal posture, the refracting angle being about 58° or 60°."

Fraunhofer confirmed Newton's views by showing that these coloured rings may be produced artificially by looking at a source of light through a plate of glass covered with fine globules of condensed vapour or with lycopodium dust. The condition necessary for the success of the experiment is that the globules should be of sensibly uniform size. He also obtained them with a large number of small metallic discs of equal size placed between two plates of glass, and he found that the diameters of the rings varied directly as the wave length, and inversely as the diameters of the discs.

M. Verdet[1] reproduced the same phenomena by covering the object-glass of a telescope with a copper plate containing a large number of small circular holes distributed irregularly. On observing a distant source of light, he saw at the focus of the telescope a system of rings similar to the coronæ.

These appearances were also observed by Young,[2] who in a very ingenious manner contrived to apply them to the measurement of the diameter of fine fibres, or small particles of any kind.

The apparatus invented by Young consisted of a metal plate perforated with a small hole of about $\frac{1}{50}$ inch (·5 mm.) in diameter. Around this aperture was a circle of smaller holes nearly half an inch in radius. The flame of a lamp was placed immediately behind the aperture, and the plate viewed through the substance under examination. The central aperture is seen surrounded by a ring, which can be brought to coincide with the circle of small holes in the plate by moving the substance backwards or forwards along a graduated scale. The distance between the substance and aperture is read off on the scale, and this varies inversely as the diameter of the ring, but theory

[1] Verdet, *Œuvres*, tom. v. p. 314.

[2] For an account of Young's Eriometer see Art. "Chromatics," vol. iii. of Supplement to *Ency. Brit.* 1817; see also *Phil. Mag.* March 1881, and vol. xx. p. 354, 1885.

shows that the diameters of the rings, produced by different sized particles, vary inversely as the diameters of the particles (see Art. 163). Consequently it follows that the diameters of the particles are directly proportional to the distances between the substance and aperture when the ring appears to coincide with the circle of small holes perforated in the plate. An experiment is therefore made with particles of a known diameter, and this gives the constant of the instrument from which the diameters of any other small particles may be determined. Thus if $d$ be the distance between the plate and substance containing particles of known radius $r$, when the ring and halo appear to coincide, and if $\delta$ and $\rho$ be the corresponding quantities for any other substance, then

$$\frac{\rho}{\delta}=\frac{r}{d}, \quad \text{or } \rho=r\frac{\delta}{d}.$$

### *Newton's Observations*

*Opticks*, book iii. part i. fourth edition, 1730.—"Grimaldo has informed us that if a beam of the sun's light be let into a dark room through a very small hole, the shadows of things in this light will be larger than they ought to be if the rays went on by the bodies in strait lines, and that these shadows have three parallel fringes, bands, or ranks of color'd light adjacent to them. But if the hole be enlarged the fringes grow broad and run into one another, so that they cannot be distinguished. These broad shadows have been reckon'd by some to proceed from the ordinary refraction of the air, but without due examination of the matter. For the circumstances of the phenomenon, so far as I have observed them, are as follows.

"*Obs.* 1.—I made in a piece of lead a small hole with a pin, whose breadth was the 42d part of an inch. For 21 of those pins laid together took up the breadth of half an inch. Through this hole I let into my darkened chamber a beam of the sun's light, and found that the shadows of hairs, thread, pins, straws, and such like slender substances placed in this beam of light were considerably broader than they ought to be if the rays of light passed on by these bodies in right lines . . . a hair of a man's head, whose breadth was but the 280th part of an inch, being held in this light, at the distance of about twelve feet from the hole, did cast a shadow which at the distance of four inches from the hair was the sixtieth part of an inch broad—that is, above four times broader than the hair, and at a distance of two feet from the hair was about the 28th part of an inch broad—that is, ten times broader than the hair, and at the distance of ten feet was the 8th part of an inch broad—that is, 35 times broader.

"Nor is it material whether the hair be encompassed with air or with any other pellucid substance. For I wetted a polished plate of glass and laid the hair in the water upon the glass, and then laying another polished plate of glass upon it, so that the water might fill up the space between the glasses, I held them in the aforesaid beam of light, so that the light might pass through them perpendicularly, and the shadow of the hair was at the same distances as big as before. . . . Therefore the great breadth of these shadows proceeds from some other cause than the refraction of the air."

He further observed that "The shadows of all bodies (metals, stones, glass, wood, horn, ice, etc.) in this light were border'd with three parallel fringes or bands of coloured light. . . . The colours proceeded in this order from the shadow : violet, indigo, pale blue, green, yellow, red; blue, yellow, red; pale blue, pale yellow, and red."

"*Obs.* 8.—I caused the edges of two knives to be ground truly strait, and pricking their points into a board so that their edges might look towards one another, and meeting near their points contain a rectilinear angle, I fastened their handles together with pitch to make this angle invariable. The distance of the edges of the knives from one another at the distance of four inches from the angular point, where the edges of the knives met, was the eighth part of an inch; and therefore the angle contained by the edges was about one degree 54′. The knives thus fixed I placed in a beam of the sun's light, let into my darken'd chamber through a hole the 42d part of an inch wide, at a distance of 10 or 15 feet from the hole, and let the light which passed between their edges fall very obliquely upon a smooth white ruler at a distance of half an inch or an inch from the knives, and there saw the fringes from the two edges of the knives run along the edges of the shadows of the

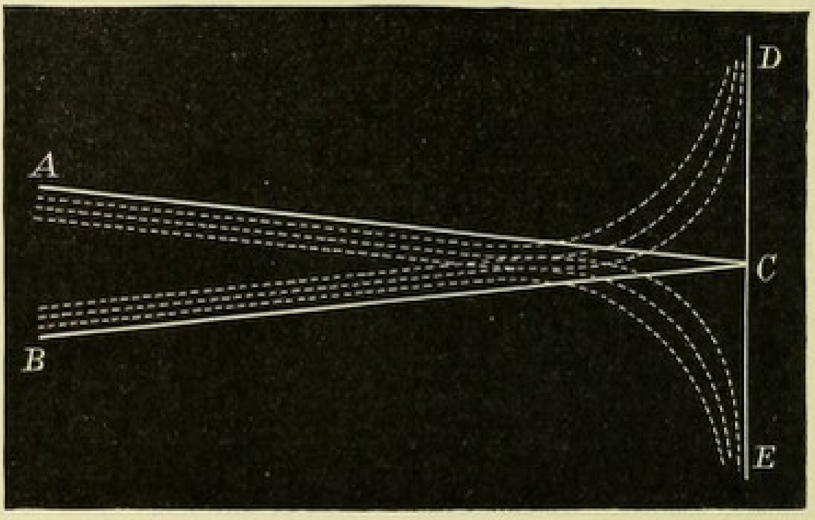



Fig. 112.

knives in lines parallel to those edges without growing sensibly broader, till they met in angles equal to the angle contained by the edges of the knives, and where they met and joined they ended without crossing one another. But if the ruler was held at a much greater distance from the knives, the fringes where they were farther from their place of their meeting were a little narrower, and became something broader and broader as they approach'd nearer and nearer to one another, and after they met they cross'd one another, and then became much broader than before.

"Whence I gather that the distances at which the fringes pass by the knives are not increased nor alter'd by the approach of the knives, but the angles in which the rays are there bent are much increased by that approach, and that the knife which is nearest any ray determines which way the ray shall be bent, and the other knife increases the bent."

"*Obs.* 10.—When the fringes of the shadows of the knives fell perpendicularly upon a paper at a great distance from the knives, they were in the form of hyperbolas. . . . Of these hyperbolas one asymptote is the line DE, and their other asymptotes are parallel to the lines CA and CB."

(In Fig. 112 CA and CB are parallel to the edges of the knives, and DE is the bisector of the external angle between them.)

Allowing the light from the small hole to pass through a prism and form a spectrum on the opposite wall, he found that the shadows of objects placed in the light between the prism and the wall were bordered with fringes of the colour of that light in which they were held. "And comparing the fringes made in the several colour'd lights, I found that those made in the red light were largest, those made in the violet light were least, and those made in the green were of a middle bigness."

## SECTION II.—DIFFRACTION GRATINGS

**135. The Diffraction Grating.**—We now proceed to the elementary explanation of the appearances presented when a distant source of light is viewed through a system of very narrow, equal, and equidistant, rectangular apertures. Such a system is named a *grating* or a *diffraction grating*. Gratings are ordinarily formed by tracing a number

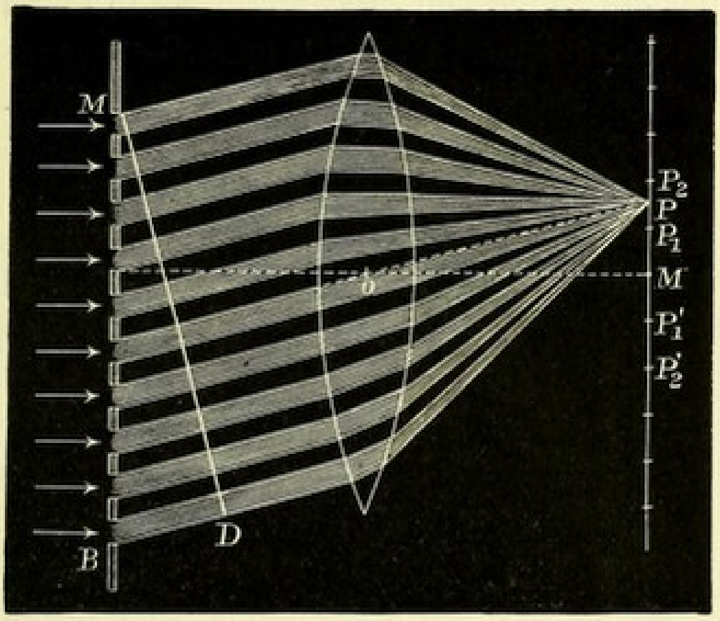



Fig. 113.

of parallel equidistant lines on a glass plate with a fine diamond point. These lines act like a system of fine opaque wires, in that the light incident on them is reflected back in all directions and refused transmission, while it passes freely through the transparent spaces between the lines. Such gratings frequently contain as many as 20,000 or even 40,000 lines to the inch, the ruling being so fine that the striæ are invisible except under a powerful microscope.

When a luminous origin is looked at through a grating a central or direct image is seen, and on either side of it there are several spectral images richly coloured with all the tints of the rainbow. These spectra increase in breadth and diminish in brilliancy as they recede from the centre, and as already remarked, they may be seen by merely viewing an ordinary candle flame or gas jet through the eyelashes, or through a pocket-handkerchief, or through a piece of ordinary wire gauze (see footnote, p. 218).

To observe the spectra to advantage a telescope should be first focussed on the luminous origin. The grating being then placed before its object-glass, the spectra are formed in its focal plane, and are viewed with all the advantages of amplification and distinctness through the eyepiece. In the following elementary explanation we shall therefore suppose the light after passing through the grating BM (Fig. 113) to fall upon a lens, which we may regard as the object-glass of a telescope.

Let $M_1N_1$, $M_2N_2$, $M_3N_3$, etc. (Fig. 114), be the apertures—that is, the transparent portions—of the grating, supposed perpendicular to the plane of the paper, and let the width of each of the apertures be $a$ while the width of each ruling is $b$, and for simplicity suppose the light to be incident perpendicularly to the grating.

Now consider the light which falls upon the lens in any direction OP (Fig. 113) where O is the optic centre of the lens. Streams of light fall upon the lens in this direction from the apertures and are brought to a focus at P, consequently the disturbance at P will be the resultant of all the disturbances sent to it by the various streams from the apertures. Draw $M_1D_1$ (Fig. 114) perpendicular to the direction OP. Each stream is propagated from this line to P in the same time. But the light is incident in the same phase at every point of the grating, since we have supposed the incident wave-front parallel to it, consequently the light which reaches P from the second aperture will be retarded on that which reaches it from the first and interference will occur. Thus the difference of path between the first element of the first aperture and the first element of the second is equal to $M_2D_2$, and the same difference exists between all the corresponding pairs of elements of these apertures, while the same remark applies to every other consecutive pair of apertures. If, therefore, $M_2D_2$ is an even number of half-wave lengths, the light from all the apertures will arrive at P in the same phase, and will reinforce each other, and the illumination at P will be very great; but if $M_2D_2$ is an odd number of half-wave lengths, the light from the first aperture will be destroyed by that from the second, the light from the third by that from the fourth, and so on. Hence in this case the illumination at P will be zero.

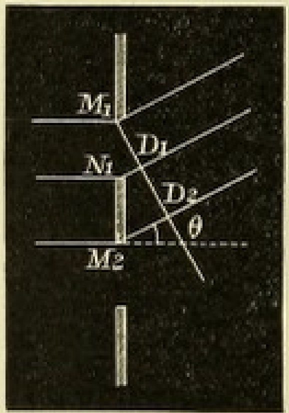


Fig. 114.

Now $M_2D_2 = (a+b)\sin\theta$ if the direction OP makes an angle $\theta$ with the normal to the grating; hence P will be very bright if

$$(a+b)\sin\theta = 2n\tfrac{1}{2}\lambda,$$

and dark if

$$(a+b)\sin\theta = (2n+1)\tfrac{1}{2}\lambda.$$

Let us now suppose $\theta$ to increase from zero while $n$ takes the consecutive values, 0, 1, 2, 3, etc. The value $\theta = 0$ corresponds to $n = 0$, so that there is no retardation, and the light from all the apertures arrives in the same phase at M on the central line OM (Fig. 113). This then is a bright point for all wave lengths, and with ordinary light will be white. An interval of darkness now occurs as $\theta$ increases from zero to the value $\theta_1$, given by the equation

$$\sin \theta_1 = \lambda/(a+b).$$

If $OP_1$ be drawn in this direction, meeting the focal plane at $P_1$, then $P_1$ is a very bright point. Another interval of darkness occurs till $\theta$ reaches the value determined by

$$\sin \theta_2 = 2\lambda/(a+b),$$

which gives another bright point $P_2$, and so on. We have therefore a succession of bright places $P_1$, $P_2$, $P_3$, etc., with dark intervals between them, and like appearances also on the lower side of the central line OM. The direction from O to the $n$th bright point is given by the equation

$$\sin \theta_n = n\lambda/(a+b).$$

White light.

What has been said so far applies to light of a definite wave length. When white light is used a brilliant rainbow-coloured band or spectrum appears at each of the points $P_1$, $P_2$, $P_3$, etc. For the angle $\theta$ corresponding to any bright point increases with the wave length, consequently the points of maximum illumination for the red light are farther removed from the centre than the corresponding points for the violet light. What were bright points at $P_1$, $P_2$, etc., with monochromatic light, are now drawn out into exquisitely coloured spectra, violet at the inner and red at the outer edge.

Over-lapping.

Several spectra are formed on each side of the central image; the first pair being separated from the second pair, and from the central image, by completely dark bands. Overlapping of the spectra will occur when the deviation of the violet of any order is less than that of the preceding red. This will take place between the spectra of order higher than the second, for since the deviation is proportional to the wave length, and since the wave length of the red is approximately twice that of the violet, it follows that the deviation of the red of the second spectrum will be approximately the same as that of the violet of the third spectrum, while the red of the third will be more deviated than the violet of the fourth, and so on. The separation of the superposed parts of the spectra at any place may be effected by means of a prism.

*Examples*

1. Explain how the lens gives only one bright point on the screen when the opaque parts of the grating are made transparent.

2. Show by Babinet's principle (Art. 133) that when a lens is used the brightness of the lateral spectra remains the same when the opaque and transparent parts of the grating are interchanged.

**136. Light incident obliquely — Minimum Deviation.** — If the incident light be not perpendicular to the plane of the grating, but falls upon it at an angle $i$ with the normal, the retardation is given by the equation

$$\delta = (a+b)(\sin\theta + \sin i).$$

For if MD and MD′ (Fig. 115) be drawn perpendicular to the incident and transmitted beams respectively, the retardation will obviously be ND + ND′; but $ND = (a+b)\sin i$ and $MD' = (a+b)\sin\theta$. Therefore, etc.

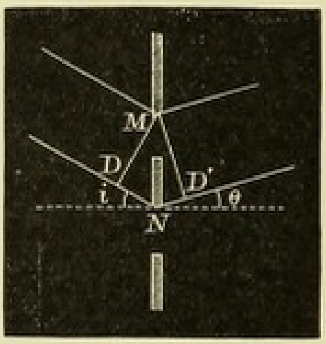



Fig. 115.

The position of the $n$th spectrum in this case is determined by the equation

$$(a+b)(\sin\theta_n + \sin i) = n\lambda.$$

The angle of incidence $i$ may be determined by measuring the angle $2i$ between the direct light and that reflected regularly from the face of the grating.

Diffraction spectra, like refraction spectra, exhibit a minimum deviation. The deviation suffered by the light of the $n$th spectrum is given by the equation

$$D = i + \theta_n.$$

But

$$(a+b)(\sin i + \sin\theta_n) = n\lambda,$$

hence the deviation of the $n$th spectrum will be a minimum [1] when $i = \theta_n$—that is, when $D = 2i$, or when the angle of incidence is equal to the angle of diffraction. We have then

$$2(a+b)\sin\tfrac{1}{2}D = n\lambda.$$

In the position of minimum deviation the definition of the spectrum

[1] For a maximum or minimum value of D we have

$$dD = di + d\theta_n = 0,$$

and from the second equation, when $\lambda$ and $n$ are given, we have

$$\cos i\,di + \cos\theta_n d\theta_n = 0.$$

Therefore $\cos i = \cos\theta_n$—that is, $i = \theta_n$, for each is less than 90°.

is considerably augmented. It has consequently been used by M. Mascart[1] in his determination of $\lambda$.

**137. Purity of the Spectra.**—We have explained how it is that when a luminous origin is viewed through a grating a series of spectra are seen on each side of the central image. We shall now show that there is no overlapping of the colours in these spectra—that, in fact, the spectra are pure. To do this let us again revert to the case of monochromatic light. With such light a series of bright images are depicted in the focal plane. If the origin be a point, we have a series of bright points; if a line parallel to the lines of the grating—such as a narrow slit—we have a series of bright lines, images of the slit.

Now we may show that the images $P_1$, $P_2$, $P_3$, etc., are really of very small dimensions, and are not drawn out or fuzzy, but are clear and well defined. For let P be a point of maximum brightness and consider light incident on the lens in a direction differing very little from OP. This light will be brought to a focus at a point very close to P, and our object is to show that at this point there is no illumination. Now since P is a bright point the retardation $M_2D_2$ is an even number of half-wave lengths, consequently the retardation $M_2D'_2$ for the new direction, very close to $M_2D_2$, is the same number of half waves plus or minus a small fraction of a wave length. Let us suppose this fraction very small, say $\frac{1}{1000}\lambda$, which will correspond to a point exceedingly close to P. The light from the first aperture is in advance of the light from the second by an amount

$$(n+\tfrac{1}{1000})\lambda,$$

consequently it is in advance of the light from the 501st aperture by 500 times this amount, or by

$$(500\,n+\tfrac{1}{2})\lambda,$$

that is, by an odd number of half waves. The light therefore from the first aperture is destroyed by the light from the 501st, the light from the second by the light from the 502nd, and so on, so that if P be a bright point there is no illumination at points even very close to it. If a narrow slit be used as the source of light, and if the lines of the grating are parallel to the slit, then for monochromatic light we will have bright lines at $P_1$, $P_2$, $P_3$, etc., parallel to the lines of the grating; but if white light be used each particular wave length gives an image of the slit, and all these are arranged side by side in a continuous spectrum without sensible overlapping or blurring.

[1] Mascart, *Ann. de l'École norm.* tom. i. et iv.; *Comptes Rendus,* tom. lvi. p. 138; tom. lviii. p. 1111.

**138. The Dispersion in the Spectrum.**—The directions to the red and violet of the $n$th spectrum being given by the equations

$$\sin\theta_r = n\lambda_r/(a+b),$$
$$\sin\theta_v = n\lambda_v/(a+b),$$

it follows that if $(a+b)$ is decreased in any proportion, then $\sin\theta_r$ and $\sin\theta_v$ will be increased in the same proportion. If these angles are small their sines are approximately equal to their tangents, so that the difference of the sines will be equal to the difference of the tangents, which will consequently be also increased in the same ratio. But the difference of the tangents is proportional to the distance between the red and violet of the spectrum, and this distance measures the amount of dispersion in the spectrum, consequently by decreasing $(a+b)$ in any ratio we increase the dispersion approximately in the same ratio. If the lines of the grating are very close and very fine, the colours of the spectrum will be widely spread out or the spectrum will be long. By this means the absence of any particular colour or wave in the solar light is exhibited, and by ruling the gratings very closely spectra have been obtained and mapped which show that the solar spectrum is not continuous, but is deficient in many places, being crossed by numerous dark lines, indicating that the corresponding wave either was not emitted by the sun, or else that it was lost by absorption or otherwise before it reached us.

Solar spectrum.

When the incident light makes an angle $i$ with the normal to the grating, and the diffracted light an angle $\theta$, we have for the $n$th spectrum $(a+b)(\sin i + \sin\theta) = n\lambda$. Hence if the angle of incidence remains constant, the variation $d\theta$ of the angle of diffraction is connected with the corresponding variation $d\lambda$ of the wave length by the equation

$$(a+b)\cos\theta d\theta = nd\lambda.$$

If $\theta$ is nearly zero the factor $(a+b)\cos\theta$ remains sensibly constant from one end of the spectrum to the other, and the variation of $\theta$ is therefore proportional to the variation of $\lambda$ only, if the angle of diffraction be small. In this case the separation, or dispersion, of the rays corresponding to the different kinds of homogeneous light is directly proportional to their difference of wave length.

When $\theta$ is not small the above equation shows that the angular variation $d\theta$ is directly proportional to the order $n$ of the spectrum, and inversely as $\cos\theta$. Consequently the higher the order of the spectrum and the greater the angle of diffraction the wider the dispersion. The dispersion, being defined as the ratio of the angular interval $d\theta$ to the corresponding variation $d\lambda$ of the wave length, is given by

$$\frac{d\theta}{d\lambda} = \frac{n}{(a+b)\cos\theta}.$$

Hence the finer the ruling of the grating the higher the dispersive power.

**139. Normal Spectrum.**—We have seen that a spectrum formed by a grating is pure,—there is no overlapping or mixture of colours. This is not generally the case with the spectra formed by glass prisms, but with proper arrangements they also may be made to give pure spectra. The grating spectra, however, have a great advantage over refraction spectra in another respect. We have seen that the dispersion depends[1] merely on the wave length and on the distance $(a+b)$ —that is, on the number of lines to the inch in the grating. Hence the spectrum formed by any grating is exactly similar to that formed by any other, one being an exact copy of the other on a larger or smaller scale,—the ratio of their lengths being $a'+b' : a+b$, for the dispersion is inversely as $a+b$. Consequently the ratio of the widths occupied in the spectrum by any two colours is invariable. This is not generally the case with the spectra formed by different prisms.[2] The relative dispersion of any two colours, the orange and blue say, may be very different with prisms made of different glasses, while some substances even reverse the order of the colours. Thus one may separate the blue and orange very much, while the other separates these colours very little. This is known as the *irrationality of dispersion*, and on account of it we are unable to compare refraction spectra, the spectrum obtained with one prism not being similar to that produced by another.

On the other hand, all spectra produced by gratings are exactly similar, and the observations made by any one at any part of the earth may be repeated and verified at any other station. The diffraction spectrum is therefore taken as the standard or *normal* spectrum.

[1] This neglects the variation of $\cos\theta$.

[2] If two prisms of different substances such as glass and water, and having refracting angles such that they give spectra of the same length, be placed with their refracting edges in opposite directions, then in a beam of light transmitted through the pair the red and violet rays will be reunited, yet the intermediate rays will be somewhat dispersed. Hence if a white line be examined through such a combination instead of being seen colourless after refraction it will form a small spectrum, purple at one end and green at the other, the water prism refracting the green or middle rays more in proportion to the extremes than the glass. These spectra formed by the irrationality of the dispersion have been called *secondary spectra* by Herschel (*Ency. Metr.*, art. "Light").

**140. Absent Spectra.**—The general conclusion which we have drawn is that if $(a+b)\sin\theta$ or $M_2D_2$ is any even number of half-wave lengths—that is, any whole number of wave lengths—then the illumination at P is a maximum, and a spectrum is formed there. We will now show that it may happen that $M_2D_2$ is an even number of half-wave lengths and yet there is no illumination at P, in fact, the spectrum is wanting. This will happen when the direction OP is such that there is an even number of half-period elements in each aperture. Each aperture will then produce zero effect at P, and there will be no illumination at that point. Now suppose that $\alpha$ and $\beta$ are the two smallest whole numbers which measure the ratio of $a$ to $b$, then $\alpha=\kappa a$ and $\beta=\kappa b$, so that if $M_2D_2=(\alpha+\beta)\lambda$ we must have $N_1D_1=\alpha\lambda$, since $M_2D_2 : N_1D_1 = a+b : a = \alpha+\beta : \alpha$. Hence there is an even number of half-wave periods in the aperture $M_1N_1$, and therefore in every aperture, consequently each aperture produces no effect at P, and the result is darkness at that point. Hence if we say that $M_2D_2=n\lambda$ corresponds to the $n$th spectrum, we may say that the

$$(\alpha+\beta)\text{th},\ 2(\alpha+\beta)\text{th, etc., } n(\alpha+\beta)\text{th}$$

spectra are wanting.

**141. Reflection Gratings.**—Spectra similar to the preceding may also be obtained by reflection, and first-class gratings may be formed by ruling very fine parallel grooves on a polished metallic surface. The streams of light regularly reflected from the polished intervals between the rulings proceed from a virtual image of the source as if they came through the intervals from behind the surface. If the surface be plane the case will be analogous to that of the transparent grating just considered, and the expression for the retardation, when the light is incident at an angle $i$ and diffracted at an angle $\theta$, becomes

$$\delta=(a+b)(\sin i\pm\sin\theta).$$

Appearances of the same nature and attributable to the same cause are often observed when a metallic surface has been polished with a rather coarse powder. The powder leaves minute striæ which affect the light as described above. A simple way of producing a similar result is by passing the finger over the surface of a piece of glass moistened with the breath. The exquisite colour of mother-of-pearl and other striated substances (formed of a vast number of very thin layers) are natural instances of the same phenomena.

**142. Curved Gratings.**—Let the surface on which the lines are ruled be not plane, but have any curved section AB (Fig. 116) perpendicular to the lines of the grating. If light from a source S fall upon it at an angle $i$ with the normal, and be diffracted at an angle $\theta$, the retardation will be

$$PQ(\sin i \pm \sin \theta).$$

Consequently for brightness we have as before, if $PQ = (a + b)$,

$$n\lambda = (a+b)(\sin i \pm \sin \theta),$$

the negative sign being taken in the case of reflection, if S and S′ lie on opposite sides of the normal (as in Fig. 116).

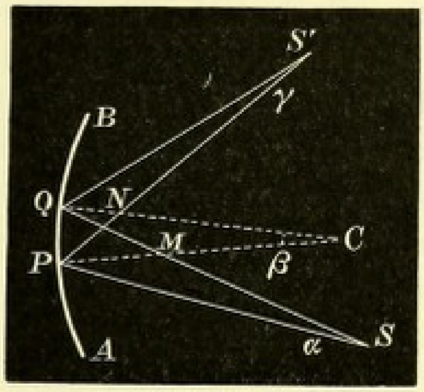


Fig. 116.

To obtain the image S′ of S, virtual or real, let us consider two rays SPS′ and SQS′ incident at angles $i$ and $i + di$, and diffracted at $\theta$ and $\theta + d\theta$ respectively, and let the normals to the curve at P and Q meet at C. Denote the small angles at S, C, and S′ by $\alpha$, $\beta$, and $\gamma$ respectively. Then clearly we have

$$\alpha + i = \beta + i + di = \text{supplement of angle at M},$$
$$\beta + \theta = \gamma + \theta + d\theta = \text{supplement of angle at N}.$$

Consequently $di = \alpha - \beta$, and $d\theta = \beta - \gamma$.

But if S′ is a focus, $\delta$ must be stationary, and this condition gives

$$\sin i - \sin \theta = \text{const.},$$

or

$$\cos i di - \cos \theta d\theta = 0.$$

Hence, substituting for $di$ and $d\theta$, we obtain

$$(\alpha - \beta)\cos i - (\beta - \gamma)\cos\theta = 0 \qquad (1).$$

Now if we write

$$PS = \rho,\ PC = R,\ PS' = \rho',\ PQ = \epsilon,$$

we have

$$\rho\alpha = \epsilon \cos i,\ R\beta = \epsilon,\ \rho'\gamma = \epsilon\cos\theta.$$

Therefore (1) becomes

$$\cos i\left(\frac{\cos i}{\rho} - \frac{1}{R}\right) - \cos\theta\left(\frac{1}{R} - \frac{\cos\theta}{\rho'}\right) = 0 \qquad (2),$$

from which we find at once

$$\rho' = \frac{R\rho\cos^2\theta}{\rho(\cos\theta + \cos i) - R\cos^2 i} \qquad (3).$$

Here we may regard $\rho$ and $i$ as the polar co-ordinates of S, and $\rho'$ and $\theta$ those of S′; hence if S describes any curve, S′ will describe another, the focal curve, defined by the above equation.

Rowland's ase.

*Cor.* 1.—If $\rho = \text{R} \cos i$, then $\rho' = \text{R} \cos \theta$—that is, if S describes a circle on R as diameter, S′ will move on the same circle. Hence if the grating curve be a circle of radius R, a source situated on a circle described on R as diameter will give spectra situated on the same circle.

This important deduction has been utilised in a masterly manner by Professor Rowland in the construction of his concave gratings.

*Cor.* 2.—When $i = \pm \theta$, if $\rho = \text{R} \cos i$ (as in Fig. 117) we have $\rho' = \rho$. Hence in the position of minimum deviation, as for prisms, the source S and its image S′ are equidistant from the grating. Consequently the diffracted rays returning to S form an image superposed on the source.

*Cor.* 3.—If the grating be plane R is infinite, and

$$\rho' = -\rho \cos^2 \theta / \cos^2 i.$$

**143. Rowland's Concave Gratings.**—Professor Rowland has successfully ruled fine gratings on a concave spherical surface of polished speculum metal. The lines are the intersections of the surface with a series of parallel equidistant planes, one of which (the central one) passes through the centre of the sphere.

Let PM (Fig. 117) be the surface of the grating, C its centre, and M the middle point of the ruled surface. On CM as diameter describe a circle. Then, as shown above (Cor. 1), a source S on this circle will give an image at S′ on the same circle, defined by the equation

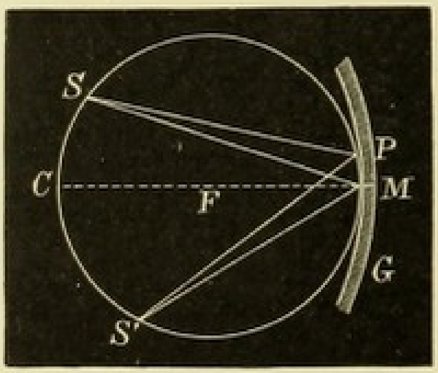


Fig. 117.

$$n\lambda = (a + b)(\sin i - \sin \theta),$$

the negative sign occurring here because we have taken S and S′ on opposite sides of the normal MC.

Real diffraction spectra will consequently be formed on the circumference of the circle SCS′, having their lines parallel to the lines of the grating. If three arms of equal length be hinged at F, the principal focus of the spherical surface, one to carry the grating G, another the slit or luminous origin S, and the third an observing telescope or screen, by rotating this latter all the spectra may be successively observed. The length of each arm is half the radius of curvature of the grating. In the third arm, instead of the screen or photographic camera, a sensitive radiometer may be substituted, and the heating effects of the various parts of the spectrum studied, as has been done by Professor S. P. Langley.[1]

[1] S. P. Langley, *Phil. Mag.* vol. xxi. p. 394, May 1886.

In order that a large part of the field of view may be in focus at once Professor Rowland places the eyepiece at C, so that $\theta = 0$, and the value of $i$ for the $n$th spectrum is then given by

$$(a+b)\sin i = n\lambda.$$

This arrangement is secured mechanically by placing the slit at S (Fig. 118), the intersection of two arms SG and SC set at right angles. At the extremities G and C of these arms the grating G and the camera or eyepiece C are placed. This arrangement is specially advantageous for photographing the spectrum. Rails are placed on SG and SC for the locomotion of the grating and camera box.

SG and SC are heavy wooden beams of which SG is fixed, while SC has a slight freedom of rotation about S, controlled by screws at C. The rails for the grating-holder and camera-box are of iron and fastened to these beams by screws which admit of adjustment, so that the rails may be straightened if the beams warp. GC is a tubular wrought-iron girder pivoted at its ends directly over the rails, on two iron carriages. Its length is approximately equal to the radius of the grating, and has a range of adjustment of about six inches. The carriages have wheels resting on the iron ways, and these enable the girder to be easily moved from one position to another. The camera-box and grating are themselves movable along GC, and have freedom of motion, but can be finally clamped in place.

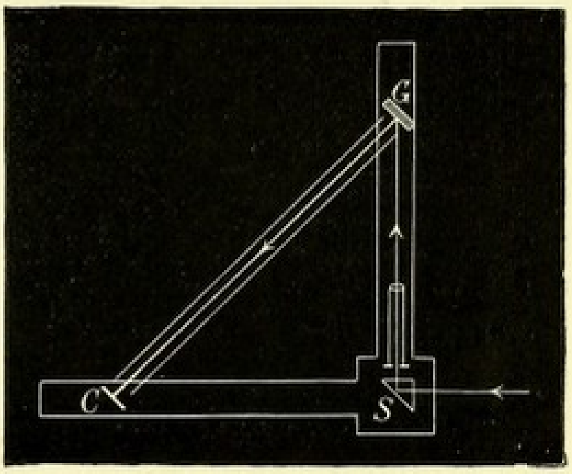


Fig. 118.—Concave Grating.

The slit, which is generally open not more than ·001 inch, can be adjusted parallel to the lines of the grating. This adjustment is one of the last to be made in mounting the grating, and is executed by turning the slit until the definition is the best possible, a condition very important in photographing the spectrum. If the slit be out $0^\circ{\cdot}5$ the definition is spoiled.

Stops can be placed at the top and bottom of the slit, thus causing the grating to be illuminated only by the centre of the solar image; otherwise the definition may be endangered by the rotation of the sun. The image of the sun on the slit should consequently be large. With the apparatus in the Johns Hopkins University it is 1·2 cm., and this is reduced one-half by the stops. For solar work the light is thrown on the slit by means of a condensing lens and a totally reflecting prism.[1] Between the lens and the prism absorbing solutions can be placed. For ordinary purposes a 10,000 grating is sufficient, but for photographing in the ultra-violet it is best to have a 20,000 grating with a ruled space of $5\frac{1}{2}$ inches on a 6-inch polished surface. The radius of curvature is generally 21·5 feet. The photographic plates are 19 inches long, 2 inches wide, and $\frac{1}{14}$ inch thick, which allows them to be bent to the required radius without breaking.

**144. Spectrum Photographs—Choice of a Grating.**—Most gratings give a brighter spectrum at one side than at the other; and so before placing the grating in the holder it must be examined to see which side should be used. Every grating has spectra of different brightness on the two sides; and one should be used which is bright in the particular spectrum desired for observation. The red of one spectrum may be bright and its violet faint. Further, the various parts of the grating, especially if it be concave, may give spectra of varying brightness. For instance, the second spectrum may be uniformly bright for all parts of the grating, while one end of the grating may give a bright third spectrum and the other a faint one. Having selected the grating which we wish to use, it is mounted in its holder and the collimating eyepiece is put in place. The focus is then carefully adjusted by altering the length of $\rho$ till the cross-hairs are exactly at the centre of curvature of the grating. On moving the bar the whole series of spectra are seen in exact focus. The rail SC on which the carriage moves is graduated to equal divisions representing wave lengths, since the wave length is proportioned to the distance SC. The instrument may thus be set to any particular wave length we desire to study, or the wave length may be obtained by a simple reading. By having a variety of scales, one for each spectrum, we can immediately see what lines are superimposed on each other, and identify them when we are measuring their relative wave lengths. Replacing the eyepiece by a camera, the spectrum may be photographed with the greatest ease. "We put in the sensitive plate either wet or dry and move to the part we wish to photograph. Having exposed that part we move to another position and expose once more. We have no thought for the focus, for that remains perfect, but simply refer to the table giving the proper exposure for that portion of the spectrum, and so have a perfect plate. Thus we can photograph the whole spectrum on one plate in a few minutes from the F line to the extreme violet, in several strips each 20 inches long, and we may photograph to the red rays by prolonged

[1] In Fig. 118 the light is thrown on the slit by means of a totally reflecting prism. This of course is necessary only when the aspect is such that the light cannot be thrown on the slit directly.

exposure. Thus the work of days with any other apparatus becomes the work of hours with this. Furthermore, each plate is to scale, an inch on any one of the strips representing *exactly* so much difference of wave length. The scales of the different orders of spectra are exactly proportional to the order. Of course the superposition of the spectra gives the relative wave lengths. To get the superposition, of course photography is the best."[1]

Between the slit and the camera-box no lens is interposed. Besides the saving of light and cost, there are no corrections necessary for spherical aberration, imperfections of lenses, etc. The concave grating is astigmatic, *i.e.* a point of light as the source is brought to focus not in a point but in a line. By the astigmatism a small spark of light at the slit is broadened out into a wide spectrum, greater accuracy in comparing solar and metallic lines is afforded, and a spectrum is obtained which is broad enough to stand enlarging.

The spectrum is normal at C. Further, in this case $n\lambda = (a+b)$ sin $i$. But $SC = \rho \sin i \propto \lambda$, thus if an absolute wave length be marked on SC and the instrument is in perfect adjustment, we can mark on the arm SC a scale of wave lengths for each spectrum, and the absolute wave length of any one line is known at once. It is important to notice that this scale on the beam is identical with the scale on the photographic plate, and that all the spectra are in focus at C at the same time, and *stay in focus;* however, C moves along SC, it being rigidly attached to G. The concave grating besides is the only spectroscope suitable for the ultra-violet and infra-red. Much longer photographic plates can be used than with any other instrument, since they can be easily bent, so as to be throughout in focus.

A 10,000 grating has on the whole better definitions than a 20,000 one, and is of course much cheaper. It is only for work with the camera in the ultra-violet part of the spectrum that it becomes necessary to use a 20,000 grating. This is due to the fact that for the same dispersion there are fewer overlapping spectra with the 20,000 than with the 10,000 grating.

A spectroscope is used for two purposes—to measure the lines in the solar or metallic spectra, or to establish coincidences simply. For both these purposes the concave grating is far superior to any other on account of the overlapping spectra.

Owing to the astigmatism of the grating it is not possible to adopt the usual method of illuminating part of the slit with the solar image

[1] Professor H. A. Rowland, *Phil. Mag.* p. 197, September 1883.

and part with the spark or arc, and so a different and far better plan is adopted. A compound photograph of the two spectra is taken. The solar spectrum is photographed along the middle of the sensitive plate; the sunlight is then turned off and the metallic spectrum is then allowed to fall upon the upper and lower part of the plate. Then, finally, the sunlight is turned on again along the middle of the plate. If there has been any gradual displacement of the camera during the operation the error is eliminated by this process if the two times of exposure to the solar spectrum are the same. A record of the barometer and thermometer readings must be kept for corrections due to variations of temperature and pressure may be considerable.

145. **Difficulties of Construction.**—The difficulties attending the construction of a grating are described by Mr. J. S. Ames:[1] "It takes months to make a perfect screw for the ruling engine, but a year may easily be spent in search of a suitable diamond point. . . . Most points make more than one 'furrow' at a time, thus giving a great deal of diffused light. Moreover, few diamond points rule with equal ease and accuracy up hill and down. This defect of unequal ruling is especially noticeable in small gratings, which should not be used for accurate work. Again, a grating never gives symmetrical spectra; and often one or two particular spectra take all the light. This is of course desirable if these bright spectra are to be used. Generally it is not so. . . . It is not easy to tell when a good ruling point is found; for a 'scratchy' grating is often a good one; and a bright ruling point always gives a 'scratchy' grating. When all goes well it takes five days and nights to rule a 6-inch grating having 20,000 lines to the inch. Comparatively no difficulty is found in ruling 14,000 lines to the inch. It is much harder to rule a glass grating than a metallic one; for to all of the above difficulties is added the one that the diamond point is continually breaking down."

146. **Measurement of Wave Lengths.**—By far the most accurate method of determining the absolute wave length corresponding to any part of the spectrum is by means of a diffraction grating. This method involves the accurate measurement of the angle of deviation of the ray under consideration, and also the measurement of the absolute length of the grating or grating space. The latter is difficult to determine accurately. Metallic gratings are much larger than glass gratings, and consequently an error in measuring them is of less importance in the result. However, it requires several days to rule a large grating, and as the coefficient of expansion of speculum metal is more than twice that of glass, changes of temperature give rise to

[1] J. S. Ames, *Phil. Mag.* May 1889.

greater irregularities in ruling, but this advantage of the glass grating is more than counterbalanced by the great difficulty in ruling one free from flaws occasioned by the breaking down of the diamond point on the hard material.

The interference methods of determining $\lambda$ usually require the exact determination of some very small length; they are therefore much inferior to the diffraction method, which lends itself more readily to linear measurement besides affording very pure spectra.

Transmission gratings may be used in two ways: (1) with the incident light perpendicular to the plane of the grating, in which case

$$n\lambda = (a+b)\sin\theta,$$

and (2) in the position of minimum deviation, when the equation which determines $\lambda$ is

$$n\lambda = 2(a+b)\sin\tfrac{1}{2}D.$$

The former method was used by Mr. Louis Bell[1] as offering fewer experimental difficulties; but with either method the accuracy with which the angular deviation can be determined far surpasses that of the measurement of the grating space $(a+b)$.

The spectrometer and grating being placed in exact adjustment, readings may be taken on the $D_1$ line in the spectra on both sides of the slit, and the angle measured five or six times in succession.

Mr. Bell worked with the third spectrum, as in it the definition was particularly good, and being of the highest order that could be conveniently observed, an error in the angle could produce little effect in the result. No correction was considered necessary for the effect of the velocity of the apparatus through space (see chap. xix.).

The accurate determination of wave lengths was first rendered possible by Fraunhofer's researches in the solar spectrum. The discovery of dark lines in the spectrum gave a definite standard of reference, and Fraunhofer[2] himself, with a wire grating, necessarily very defective, obtained very fair determinations of the wave length of the D line.

His mean result was ·0005888 mm., which is remarkably accurate, considering his gratings and the fact that most of his angles of deviation were less than 1°.

These determinations were not improved on till the importance of

[1] Louis Bell, *Phil. Mag.* p. 265, March 1887.

[2] "Neue Modifikation des Lichtes durch gegenseitige Einwirkung und Beugung der Strahlen, und Gesetze derselben," presented to the Munich Academy in 1821. See Bell, "On the Absolute Wave Length of Light," *Phil. Mag.* p. 246, April 1888.

spectroscopic work was established by the great researches of Bunsen and Kirchhoff, and the art of ruling gratings was much improved by Nobert. Mascart employed four or five of Nobert's gratings, and worked with them in the position of minimum deviation—that is, so that the incident and diffracted rays made equal angles with the plane of the grating. This method avoids the necessity of placing the grating perpendicular to the axis of either telescope, but it is rather more difficult in the experimental work, and is perhaps of questionable utility. However, it generally improves the definition, and is capable of giving very accurate results.

In the same year (1868) Ångström's researches appeared, and for long remained the standard of reference in all questions of wave length. He used Nobert's gratings, and in spite of the fact that these were small and inaccurately ruled, giving imperfect definition and showing numerous "ghosts," his results would have been very nearly exact if his standards of length had been correct.

In all the earlier determinations of the wave length insufficient attention was paid to the measurement of the grating spaces, and this of course requires an accurate standard of length. Thalén,[1] who assisted Ångström in his work, corrected it afterwards for the error in the assumed length of the Upsala metre.

Ten years after Ångström's research Mr. C. S. Peirce[2] again attacked the problem with Rutherford gratings far superior to any previously used.

The grating space $(a+b)$ is never perfectly uniform throughout the whole extent of the ruled surface. Regular or periodic variations produce "ghosts" and differences in focus of the spectra on opposite sides of the centre. There are other variations of an irregular character, such as the displacement or omission of one or more lines, or, what is far worse, the more or less sudden change in the grating space, forming a part having a grating space peculiar to itself. This latter is by far the most formidable type of error. The other irregularities are harmless, and occur in most gratings. If the abnormal portion of the grating be confined to a few hundred lines, they will merely diffuse a certain amount of light without producing false lines or sensibly injuring the definition. They will, however, lead to an incorrect result if we determine the grating space by measuring the total length of the grating, and dividing by the total number of lines.

Mr. Bell[3] describes an experiment illustrating the effect of these

[1] Thalén, *Sur le Spectre du Fer*, Upsala, 1885.
[2] *American Journal of Science*, third series, xviii. p. 51, 1879.
[3] Bell, *Phil. Mag.* p. 362, May 1888.

errors of ruling: "Place a rather bad grating—unfortunately only too easily obtained—on the spectrometer, and, setting the cross-hairs carefully on a prominent line, gradually cover the grating with a bit of paper, slowly moving it along from one end. In very few cases will the line stay upon the cross-hairs. A typical succession of changes in the spectrum is as follows:—Perhaps no change is observed until two-thirds of the grating has been covered. Then a faint shading appears on one side of the line, grows stronger as more and more of the grating is covered, and finally is terminated by a faint line. Then this line grows stronger till the original line appears double and finally disappears, leaving the displaced line due to the abnormal grating space."

The effect of an abnormal portion of the grating is, therefore, to cause a displacement of the lines of the spectrum and lead to error in the evaluation of the deviation, as well as in the calculation of the grating interval. The abnormal spacing generally occurs at the end of the grating where the ruling was begun, for the engine after starting requires some little time to settle down to a uniform state.

To detect and evaluate the errors of irregular spacing Mr. Bell[1] proposes the *calibration* of the grating. In this process the grating is examined under a microscope from end to end. This gives the variations in the lengths of the spaces in different parts of the grating. Bell's mean result for $D_1$ with four corrected gratings, two on glass and two on speculum metal, gave at 20° C. and 760 mm. pressure in air

$$\lambda_{D_1} = 5896{\cdot}18 \text{ tenth metres,}$$

or in vacuo

$$\lambda_{D_1} = 5897{\cdot}90 \text{ tenth metres,}$$

which, as far as errors of observation go, he considers should be correct to within one part in half a million. The wave lengths of the other lines in the spectrum derived from this are in air at 20° and 760 mm.

| | | | |
|---|---|---|---|
| A (line between "head" and "tail" of group) | 7621·31 | $E_1$ (line between "head" and "tail" of group) | 5270·52 |
| B | 6884·11 | $E_2$ | 5269·84 |
| C | 6563·07 | $C_1$ | 5183·82 |
| $D_1$ | 5896·18 | F | 4861·51 |
| $D_2$ | 5890·22 | | |

Mr. Bell has given with the above figures the chief results previously obtained for $D_1$ as follows:—

[1] Louis Bell, *Phil. Mag.* p. 363, May 1889.

| | | | |
|---|---|---|---|
| Mascart | 5894·3 | Ångström corrected by Thalén | 5895·89 |
| Van der Willigen | 5898·6 | Müller and Kempf | 5896·25 |
| Ångström | 5895·13 | Macé de Lépinay | 5896·04 |
| Ditscheiner | 5897·4 | Kurlbaum | 5895·90 |
| Peirce | 5896·27 | Bell | 5896·18 |

Every method of determining wave lengths must necessarily involve the uncertainties of the standards of length used. The experimental difficulties involve small but troublesome corrections, such as the effect of moisture in the atmosphere, changes of pressure, uncertainty as to the true temperature of the grating, and variations of the grating space.

[Runge (*Astronomy and Astrophysics*, vol. xii. p. 426) gives an example on the necessity of allowing for the dispersion of air in determining wave lengths by the method of coincidences: "Suppose at a temperature of 20° Celsius and a pressure of 760 mm. two rays have the wave lengths 6000 and 2000. They would then exactly coincide if one was observed in the first and the other in the third order. Now reduce both rays to vacuo. They would no longer coincide, the ray in the third order lying to the less refrangible side by $0{\cdot}635 \times 3 - 1{\cdot}633 = 0{\cdot}272$. Since the corrections are proportional to the density of the air a barometric pressure of say 770 mm. and a temperature of 12° Celsius would make the distance between the two lines 0·011, an amount which is larger than the probable error that may be reached with well-defined lines."

He gives the following table for reducing to vacuo Rowland's standard wave lengths and all wave lengths that are determined from them by interpolation:—

| Wave Lengths. | For Reduction to Vacuo add | Wave Lengths. | For Reduction to Vacuo add |
|---|---|---|---|
| 8000 | 2·164 Ångström | 4000 | 1·109 Ångström |
| 7000 | 1·893 ,, | 3000 | 0·857 ,, |
| 6000 | 1·633 ,, | 2500 | 0·739 ,, |
| 5000 | 1·369 ,, | 2000 | 0·635 ,, |

Compare the table on p. 137 containing Kayser and Runge's indices for dry air.]

[**146*a*. The Échelon Spectroscope.**[1]—We have seen (Art. 138) that the dispersion of a grating varies directly as the order ($n$) of the spectrum and inversely as the distance ($a + b$) between consecutive rulings.

[1] It is sad to recall that Prof. Preston had ordered an échelon spectroscope for his investigations on the Zeeman effect, and that he had died before the instrument was completed.

Consequently to increase the dispersion the rulings must be made finer or else a spectrum of higher order must be observed. On account of the extreme faintness of high order spectra the second alternative afforded no promise of improvement in the construction of gratings until Michelson[1] investigated the subject. He pointed out the conditions depending on the shape of the grooves made in ruling under which a reflecting grating would concentrate most of the light into a particular group of spectra, and he constructed a most ingenious and valuable transmission grating by means of which the same effect is produced. Michelson took a number of plane-parallel plates of optical glass of exactly the same thickness ($e$), and built them up in the form of a stair, each step being of the same depth ($f$).

Let $PACQ$ and $P_1B_1A_1Q_1$ represent rays of light normally incident on the plates and passing through corresponding points $A$ and $A_1$ on consecutive steps. Draw $A_1C$ at right angles to $CQ$. The retardation of the second ray on the first is $\mu B_1A_1 - AC$ if $\mu$ is the refractive index of the glass. Now $AC$ is equal to the sum of the projections of $AB_1$, $B_1A_1$, and $A_1C$ on its direction, that is $AC = -AB_1 \sin\theta + B_1A_1\cos\theta = -f\sin\theta + e\cos\theta$. The retardation is therefore

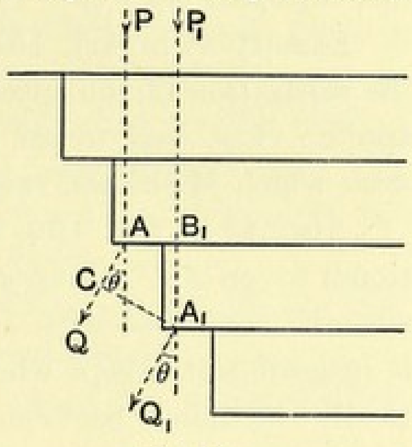



Fig. 118*a*.

$$m\lambda = \mu e + f\sin\theta - e\cos\theta \quad . \quad . \quad . \quad . \quad . \quad (1)$$

and if $m$ is a whole number the emergent light of wave length $\lambda$ from all the steps making the angle $\theta$ with the normal will form a bright image of the slit when brought to a focus by a lens. As the retardation is large, $m$ will be large and the order of the spectrum will be very high.

Differentiating (1) for a given value of $m$ we find

$$d\theta = \frac{m\,d\lambda - e\,d\mu}{e\sin\theta + f\cos\theta} \quad . \quad . \quad . \quad . \quad . \quad . \quad (2)$$

and also replacing $m$ by $m+1$ and $\theta$ by $\theta + \Delta\theta$ in (1) we have

$$\Delta\theta = \frac{\lambda}{e\sin\theta + f\cos\theta} \quad . \quad . \quad . \quad . \quad . \quad . \quad (3)$$

so that on division

$$\frac{d\theta}{\Delta\theta} = \frac{m\,d\lambda - e\,d\mu}{\lambda} \quad . \quad . \quad . \quad . \quad . \quad . \quad (4)$$

where $d\theta$ is the change of $\theta$ due to change of wave length ($d\lambda$), and $\Delta\theta$ is its change on passing to the spectrum of next highest order ($m+1$).

[1] A. A. Michelson, *The Astrophysical Journal*, vol. viii. p. 37 (June 1898).

Replacing $m$ by its value from (1) we have

$$\frac{d\theta}{\Delta\theta} = \frac{f \sin\theta - e\cos\theta}{\lambda}\frac{d\lambda}{\lambda} - ed \cdot \frac{\mu}{\lambda} \quad . \quad . \quad . \quad . \quad . \quad (5)$$

For flint-glass $d(\mu/\lambda)$ is approximately equal to $-2d\lambda/\lambda^2$, and the formula becomes

$$\frac{d\theta}{\Delta\theta} = \frac{f\sin\theta + e(2-\cos\theta)}{\lambda} \cdot \frac{d\lambda}{\lambda} \cdot = \left(2 - \frac{\sin(\alpha-\theta)}{\sin\alpha}\right)\frac{e}{\lambda} \cdot \frac{d\lambda}{\lambda} \quad . \quad . \quad . \quad (6)$$

where $\alpha(=\tan^{-1}e/f)$ is the angle of slope of the échelon. If $d\lambda/\lambda = 0{\cdot}001$ as in the case of the two sodium lines, if $e = 5$ mm. $= 10000\lambda$, and if $\theta$ is very small, $d\theta = 10\Delta\theta$ approximately, or the sodium lines would be separated by ten times the distance between the spectra. In this case also $m\lambda = (\mu - 1)e = 10000(\mu - 1)\lambda$, so $m$ is about 5000.

Exactly as in Art. 157 when the échelon is composed of $n$ plates the separation of the spectra ($\Delta\theta$) is increased $n$-fold. Indeed, Fig. 140 applies very well to an instrument composed of only *seven* elements with which Michelson easily observed the Zeeman effect.

Also, as in Art. 156, the intensity due to a single step is proportional to $\sin^2\phi/\phi^2$ where $\phi = \pi f\lambda^{-1}\sin\theta$. On account of the enormous size of $f$ compared with $\lambda$ the total intensity (compare again Art. 156) is insignificant except where $\theta$ is very small, and all the light is practically included between the deviations $\pm\lambda/f$. But the distance between two successive spectra is $\lambda/f$ by (3), so there will in general be two spectra visible. By slightly inclining the échelon one of these may be rendered central and much more intense than the other.

"The overlapping of the spectra is overcome by a direct vision prism of moderate dispersion, but the distance between the spectra is so small in comparison with the dispersion of the échelon that the spectra of the source under examination must consist of rather fine lines if overlapping is to be avoided."

Michelson gives the dimensions of a twenty plate échelon constructed for him. Each plate was 18 mm. thick, and the successive elements diminished in width from 22 mm. to 2 mm., so that the width of the elementary pencils is 1 mm., and the retardations are of the order of 20,000 waves. The resolving power of this instrument is about three times that of the best gratings, and more powerful instruments can be constructed. The real difficulty is to obtain plates of exactly the same thickness. Most valuable results may be anticipated in the analysis of close groups of lines.]

## Section III.—Investigation of the Intensity in Diffraction Patterns

**147. Application of the Graphic Method.**—In the foregoing sections we have considered merely the general character of the effects produced by diffraction when light, diverging from a luminous point, falls upon a narrow aperture or passes by the edge of an opaque obstacle. The actual calculation of the intensity of the illumination at any point of a diffraction pattern (that is, the fringe system produced by diffraction) is generally a problem of some difficulty when attacked directly by the analytical method, but in many cases the solution may be effected with great simplicity by means of the elegant graphic method introduced in Art. 45. We shall therefore recapitulate briefly this method of representing the resultant of a system of vibrations of the same period, but of different amplitudes and phases.

Thus it has been shown that if a polygon be constructed so that the lengths of its sides (Fig. 119) represent the amplitudes $a_1, a_2, \ldots a_n$ of a system of vibrations simultaneously superposed on a particle, while the angles which these sides make with a given line OX represent the phases of the corresponding vibrations at any instant, then the closing side OP of the polygon represents the amplitude, and the angle XOP which it makes with OX represents the phase of the resultant vibration. In this figure the phases of the successive vibrations are taken in ascending order of magnitude, and are such that there is an abrupt change in passing from each to those adjacent to it. If, however, the change of phase be not abrupt, but varies continuously in passing from each to the next, and if the amplitudes be very small, then the sides of the polygon will be very short, and the angles which they make with each other will be very small, so that in the limit the figure becomes a continuous curve, as shown in Fig. 120. This is what happens when we attempt to represent the effect of a complete wave, or of any part of it, at an external point.

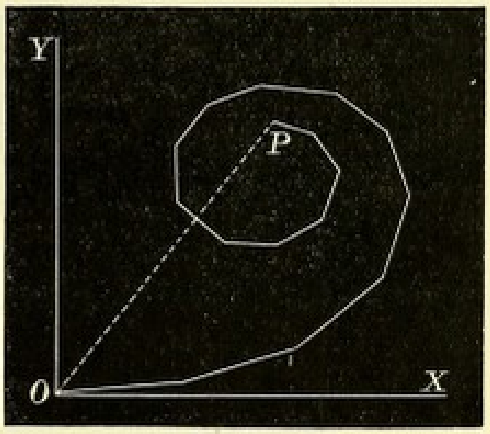


Fig. 119.

The shape of such a curve depends on the manner in which the amplitudes and phases of the vibrations change in passing from each to the next of the system, the curvature at any point being measured by the rate of variation of phase with amplitude. Thus if the element of length of the curve be denoted by $ds$, and if the angle between two consecutive elements be $d\phi$, then $d\phi$ will be the difference of phase between two consecutive vibrations, and the radius of curvature at the corresponding point of the vibration curve will be $ds/d\phi$. The form of the curve representing the effect of any part of a wave will consequently depend on the manner in which the wave is subdivided into elements of area.

Curvature.

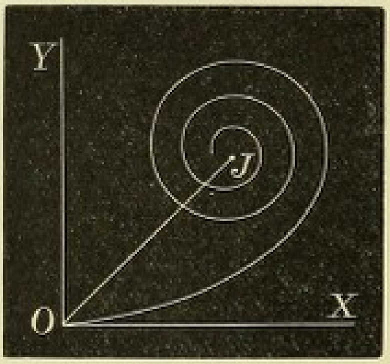



Fig. 120.

In the particular case in which the vibrations form a system of equal amplitude and uniformly increasing phase the curvature is the same at all points, and the vibration curve (as already noticed in Art. 45) is a circle. This is the case discussed in Arts. 51 and 52, in which the influence of obliquity is neglected, and where the wave front is divided into ring elements corresponding to equal increments of phase. The areas of these elements are not equal, but vary directly as their distances from the point O, at which their joint effect is to be calculated, and it follows therefore that when the influence of distance only is considered, the amplitudes of the vibrations produced at O by the various rings are the same. When the influence of obliquity is taken into account, on the other hand, the amplitudes form a diminishing series, and therefore in the vibration curve $ds$ continually diminishes

The ring method.

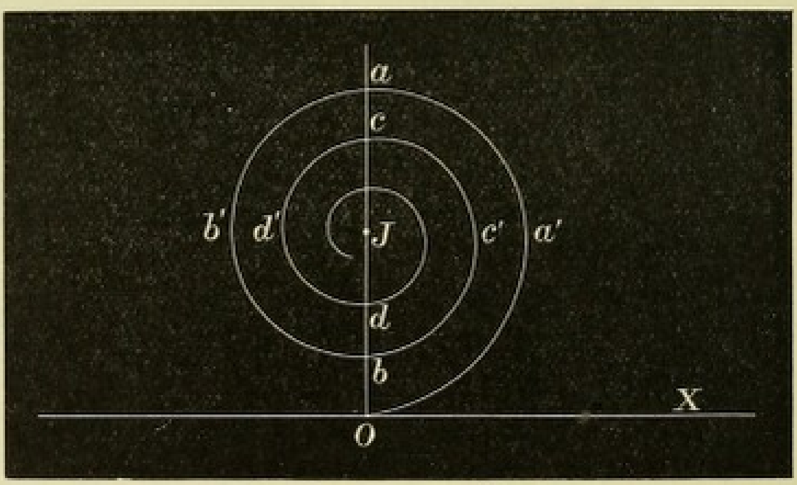



Fig. 121.—Vibration Spiral.

as $\phi$ uniformly increases, so that the curvature gradually increases as we proceed along the curve, and consequently the effect of the wave, instead of being represented by a circle, is represented by a spiral curve encircling a point J (Fig. 121) with convolutions of ever diminishing radius.

This method of dividing the wave front into elements of equally increasing phase is undoubtedly the most desirable when it can be applied, but there are many cases in which it ceases to be convenient, and other methods of subdivision have to be adopted. For example, the ring method applies at once if we wish to calculate the intensity at any point on the axis of a circular aperture, or the effect of any circular or annular portion of a wave at a point on its axis. When the aperture is rectangular, or when the light is diffracted over a simple straight edge, it is best to divide the wave into elementary strips parallel to the edge of the obstacle. We shall consequently consider this method of strip division in some detail before applying it to particular problems. In the ring method the surface of the wave is divided into elements of area by means of a system of spheres having a common centre, while in the strip method the surface is intersected and divided into strip areas by means of a system of planes having a common edge. The strip method.

**148. The Method of Strip Division.**—Let two parallel right lines AB and CD (Fig. 122) be drawn in the front of a plane wave so as to include a very narrow strip of the wave, and let it be required to determine the effect of this strip at any external point O. Taking the strip to be very narrow we may represent it by the right line AB (Fig. 123), and divide it into elements of length, $PM_1$, $M_1M_2$, etc., by points taken on it at equally increasing distances from O. These elements are such that

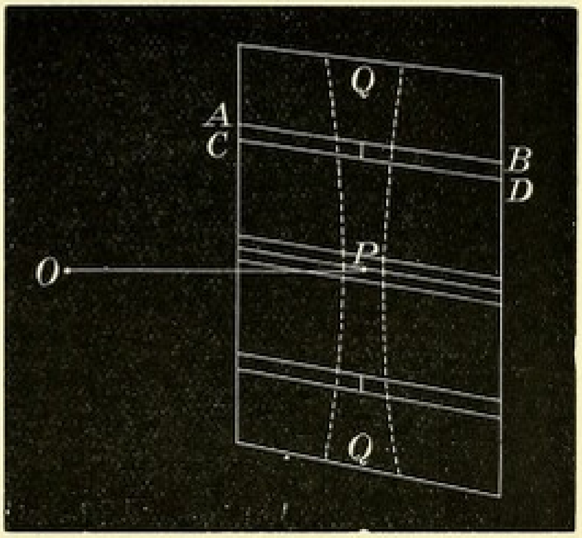


Fig. 122.

$$OM_1 - OP = OM_2 - OM_1 = \text{etc.} = OX' - OX = \delta.$$

They are elements of uniformly increasing phase, but are not of equal length. To express the length of any one of them, XX′ for example, we have by similar triangles

$$(OX' - OX) : XX' :: PX' : OX'.$$

Denoting OX′ − OX by $\delta$, and writing $r$ and $x$ for OX′ and PX′—that is, for the distances of the element from O and P respectively—we have [1]

$$XX' = \frac{r\delta}{x}.$$

Hence, if we neglect the influence of obliquity, the amplitude of the vibration contributed by the element XX′ will be proportional to

$$\frac{XX'}{r} = \frac{\delta}{x}.$$

We conclude therefore that when the strip is divided into elements corresponding to uniformly increasing phases their effects diminish as they recede from the pole P, the diminution of effect being inversely as the distance of the element from the pole of the strip. When the influence of obliquity is taken into account this rate of decrease in the effects of the elements, as they recede from the pole, is made still more rapid. Hence when we compare two consecutive half-period elements of the strip, as in Art. 51, it follows that when they are near the pole they differ considerably in effect, but when they are far away from the pole they approximately neutralise each other.

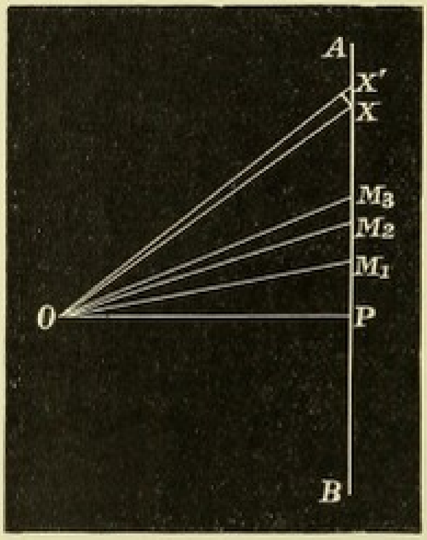


Fig. 123.

Denoting the effects of the consecutive half-period elements by $m_1$, $m_2$, $m_3$, etc., and taking into account the two halves of the strip, viz. AP and BP, the whole effect of the strip may be written in the form

Single strip.

$$S = 2(m_1 - m_2 + m_3 - m_4 + \text{etc.}).$$

In this series the terms diminish rapidly at first, so that they soon become very small, and ultimately equal and opposite. We may therefore conclude that when the wave length is small the whole effective portion of the strip is confined to a small region in the neighbourhood of the pole. In the same way every other strip may be reduced to a small effective portion in the neighbourhood of its pole with respect to O, and the whole wave may be replaced by a narrow equatorial band QQ (Fig. 122) passing through P in a direction at right angles to the strips.

The calculation of the effect of the whole wave at O is consequently reduced to that of the equatorial band, QQ. Now this band

[1] This merely expresses that $rdr = xdx$.

is cut into elements by the strips ABCD, etc., which correspond to equal increments of phase, and as the effects of these at O vary inversely as their distances from P (when the effect of obliquity is neglected), it follows that the successive elements produce effects at O which rapidly diminish as they recede from the pole. This diminution is rendered still more rapid by the influence of obliquity, which also causes a falling off in amplitude as we recede from the pole. The result is that the effective portion of the equatorial band is confined to a small area in the neighbourhood of P, and, as before, the effect of the whole wave may be limited to a small area surrounding the pole. Equatorial band.

Now if the effect of a single strip, such as AB (Fig. 123), be represented graphically in the manner already explained, it is clear that (since the amplitudes of the vibrations contributed by the successive elements of the strip rapidly diminish as they recede from the pole) the curvature of the vibration curve will increase rapidly at first, and then more slowly, so that the curve will be a spiral similar to that shown in Fig. 120, with convolutions of ever decreasing radius encircling a point J. This spiral represents one half of the strip, and the other half will be represented by the same spiral repeated. The effect of the whole strip will consequently be represented in amplitude by 2OJ, and in phase by the angle XOJ. This angle measures the difference in phase between the resultant vibration contributed by the whole strip, and that contributed by the central point or pole of the strip. Single strip.

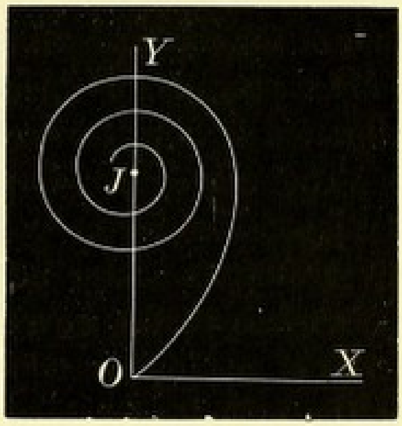


Fig. 124.

Having determined the amplitude and phase of the vibration contributed by each strip — that is, by each element of the equatorial band QQ—a spiral may be drawn in the same way to represent the effect of the complete wave. Now since the phase of the vibration contributed by the whole wave differs from that arriving from the pole by 90° (Art. 52), it follows that the spiral representing the equatorial band must encircle a point J on the axis OY (Fig. 124), and must start from O, not as a tangent to OX, but, making an angle with it, determined by phase difference between the resultant vibration of a whole strip and that contributed by its pole. This spiral represents one half of the equatorial band, and the other half is represented by the same spiral repeated, so that the whole amplitude is measured by 2OJ. Whole wave.

In the case of a spherical wave the surface may be divided into a series of strips by a system of planes passing through a diameter of

Spherical wave.

the sphere. Thus if APB (Fig. 125) represents the front of a spherical wave diverging from C, and if it is required to calculate the effect at O of the whole, or some portion of it, the surface of the wave may be divided into a system of strip areas by a system of planes passing through a diameter AB of the wave drawn at right angles to OC. These planes, which may be called meridian planes, intersect the surface in great circles, and divide it into elementary strips, which are lunes of small area. Each of these strips may be reduced in effect to a small area in the neighbourhood of its pole, and may be represented by a spiral curve after the manner of Fig. 120. It follows therefore that, as in the case of a plane wave, the spherical wave may be replaced in effect by an equatorial band lying along the great circle PMN, which cuts the meridian planes orthogonally. This band in turn may be represented graphically by a spiral curve after the manner of Fig. 124, and its effective portion is restricted to a small area in the neighbourhood of P. When represented in this manner the amplitude of the resultant vibration is 2OJ (Fig. 124), and its phase is 90° in advance of that of the vibration arriving from the pole of the wave.

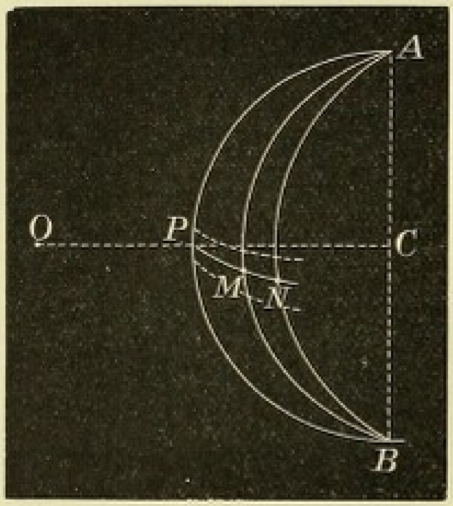


Fig. 125.

Cylindrical wave.

The only other form of wave which we need consider is the cylindrical wave, such as diverges from a long narrow slit. In this case the surface of the wave may be divided into a series of rectilinear strips parallel to the length of the cylinder by a system of planes drawn through its axis, as shown in Fig. 126. Each of these strips may be reduced to a small effective portion, and the whole wave as before may be replaced by an equatorial band PMN. This in turn reduces to a small portion around P, and may be represented graphically, as in the case of a plane wave.

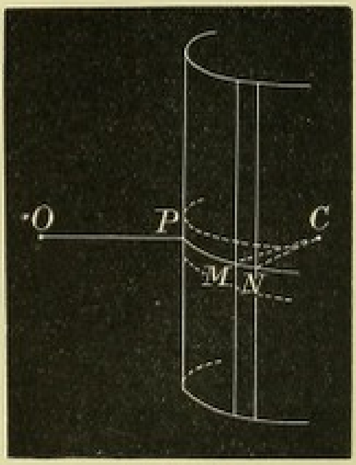


Fig. 126.

**149. First Applications of the Spiral.** — Before making any calculation of the actual intensity at the various points of a diffraction pattern, we shall consider in a general manner the fluctuations of intensity in some of the elementary cases already noticed[1] (Sec. I.).

[1] This application was published by M. Cornu as a "Méthode nouvelle pour la discussion des problèmes de diffraction dans le cas d'une onde cylindrique" (*Journal de Physique*, tom. iii. 1874).

For this purpose, when the strip method of division is employed, each half of the wave gives rise to a spiral curve, and these spirals are identical in shape and position. Instead of superposing them, however, it will often be found convenient, for the purposes of representation, to draw one of them above the line OX, and the other below it, as shown in Fig. 127. The chord joining any two points on the upper spiral will represent the resultant effect of a corresponding portion of one half of the wave, and the chord joining two points on the lower spiral will represent a portion of the other half of the wave. Thus OJ represents the whole effect of the upper half of the wave, and OM represents the effect of a portion of the lower, whereas JM is the resultant of these two (viz. OJ and OM) taken together. This applies to the case of light diffracted over a straight edge.

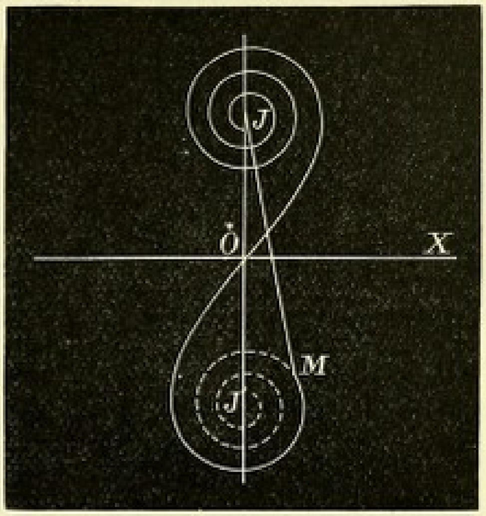


Fig. 127.

*Straight Edge.*—It has been already indicated that when light passes by the edge of an opaque screen a system of fringes exists just outside the geometrical shadow, but that inside the shadow the light fades away gradually without passing through any alternations of brightness and darkness. The illumination at any point outside the shadow is contributed by a complete half wave RS (Fig. 107) and by fraction RA. The effect of the half wave will be represented by OJ and the fraction by OM (Fig. 127), where M is some point on the other half of the spiral. The resultant effect will therefore be represented by JM. Hence as M moves along the spiral JM passes through a series of maxima and minima. There is consequently a series of alternations of brightness and darkness outside the geometrical shadow. For as the point on the screen moves outwards from the shadow, the point M moves round on the spiral towards J′, and JM passes through a maximum and a minimum every convolution. The least value is JO, that due to half a wave. The intensity then rises and falls, the values of the maxima being greater than JJ′, that due to the complete wave, and the values of the minima greater than JO, that due to half a wave. At the edge of the geometrical shadow the amplitude is OJ, and the intensity is consequently one-fourth of that produced by the whole wave.

For the illumination at any point within the geometrical shadow we have only to deal with part of the spiral OJ (Fig. 128), as only a fraction of half a wave is now in action. As P recedes within the shadow the tracing point M recedes from O along the spiral OJ, moving round towards the point J. The line JM, which represents the resultant vibration, continually decreases towards zero, without passing through any maximum or minimum values. The illumination therefore falls off gradually to zero within the shadow.

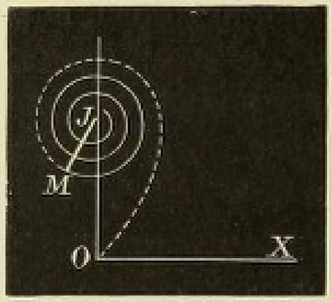



Fig. 128.

*Narrow Aperture.*—Let us now consider the illumination produced at any point by a very narrow rectangular slit. Since the element of arc of the spiral measures the amplitude of vibration of a corresponding element of the wave, it follows that the length of the complete arc of the spiral which represents the wave from the slit is simply proportional to the width of the slit; consequently the amplitude of the resultant vibration at any point on the screen will be measured by the right line joining the extremities of a constant length of the curve, viz. an arc proportional to the width of the slit. Inside the geometrical projection the arc in question passes through O and belongs partly to one half of the spiral and partly to the other. Outside the projection the arc is situated altogether in one half of the spiral. In all cases it is clear that there will be generally fluctuations of intensity at different points of the screen.

If the slit be very narrow, so that the corresponding arc of the spiral is small, then inside the projection the intensity will remain constant over a considerable range, and be very nearly proportional to the square of the width of the aperture, since here the arc will nearly agree with its chord.

*Narrow Wire.*—The case of a narrow wire can be deduced from that of a narrow rectangular aperture of the same dimensions. Inside the shadow of the wire the effect of that part of the wave which passes one side of the wire is represented by JM (Fig. 129) where M is some point on the spiral OJ, and the effect of the part passing the other side of the wire is represented by J′M′ where M′ is some point in the spiral OJ′, and the arc MM′ is proportional to the width of the wire. If the wire subtends a considerable number of half-period elements at the screen, the arcs

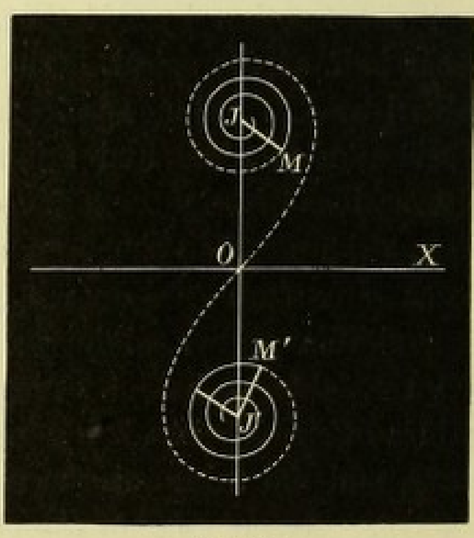



Fig. 129.

OM and OM′ will contain several convolutions of the spirals and the lines JM and J′M′ will be nearly equal, so that there will be destructive interference when JM and J′M′ are in opposite directions, and maximum illuminations when they are in the same direction. In the interior of the shadow we have thus a system of interference bands. As we approach the borders of the shadow, the point M, suppose, moves towards O on the spiral and the point M′ moves towards J′, since the arc MM′ must be of constant length. The lines JM and J′M′ in this case differ considerably in magnitude, so that when they are parallel and in opposite directions there will still be some resultant illumination and the minima will not be places of complete darkness. Outside the shadow we have a complete half wave and a fraction from one side of the wire. These will be represented by OJ and OM′ respectively, while from the other side we have a portion represented by J′M″, where M″ is some point on the spiral near J′. The arc M′M″ is absent and of a constant length proportional to the width of the wire. Thus if from J′ we draw J′N (Fig. 130) parallel and equal to the chord of the arc M′M″, then since JJ′ represents the effect of the whole wave, and since J′N represents the effect of the intercepted portion, it follows that JN will represent the transmitted portion. JN therefore represents the effect at a point outside the shadow, and as J′N revolves round J′—that is, as the point recedes from the edge of the shadow—the line JN will pass through a series of maxima and minima, which represent the external fringes.

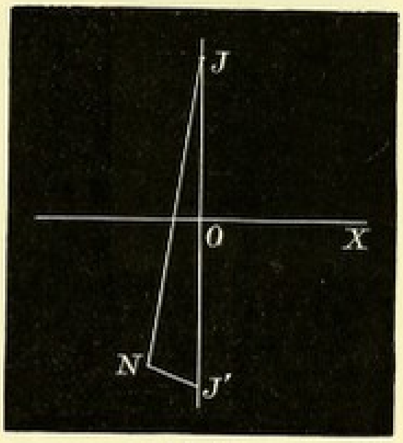


Fig. 130.

In the case of the narrow aperture it is JN that is intercepted and J′N transmitted.

*Two Narrow Rectangular Apertures.*—In the case of two equal narrow apertures we have to consider the resultant of two arcs of the spiral of lengths proportional to the widths of the apertures. These arcs are separated by an arc of constant length, proportional to the distance between the apertures. The resultant is therefore the vector sum of the chords of the two arcs—that is, the diagonal of the parallelogram having its adjacent sides parallel and equal to the chords. This diagonal will pass through a maximum or minimum value according as the chords of the arcs are parallel and in the same or opposite directions.

Two narrow wires give a corresponding system of fringes, and the diffraction patterns afforded by other arrangements may be investigated in a similar manner.

### *Calculation of the Intensity in the Case of Parallel Light*

**150. Narrow Rectangular Aperture.**—It is very easy to examine by the graphic method the case in which the incident light is a parallel beam, or in other words, a plane wave. We shall first take the case of a narrow rectangular aperture of width $a$, and at present we shall omit all consideration of the length of the aperture and investigate only those phenomena which arise from the narrowness of its width.

Let AB (Fig. 131) be a cross section of the aperture by a plane drawn at right angles to its length, and let the incident light make an angle $90^\circ - i$ with the width AB; it is required to determine the intensity of the illumination at any point on the other side of the aperture.[1] For this purpose let us take the beam of diffracted light which leaves the slit in any given direction AX, making[2] an angle $90^\circ - \theta$ with AB. Then it follows, as in Art. 136, that the relative path retardation of the extreme rays AX and BY is

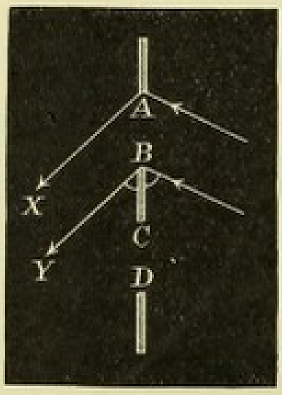


Fig. 131.

$$\delta = a\,(\sin i + \sin \theta),$$

and their difference of phase is found by multiplying this by $2\pi/\lambda$. Now let the aperture be divided into a very great number of exceedingly narrow strips of equal width, the lengths of the strips being parallel to the length of the aperture. This is equivalent to dividing AB into a great number of small elements of equal length, and since the light is parallel it is clear that each of these elements produces vibrations of equal amplitude and corresponds to equal increments of phase. The problem therefore reduces to the calculation of the resultant of a number of vibrations of equal amplitudes and uniformly increasing phases. Such a system, we have already seen (Art. 45), gives a vibration curve of uniform curvature—that is, a circle—and the resultant is consequently represented in amplitude and phase by the length and direction of a chord OM (Fig. 132) of a circle. The

[1] For the sake of definiteness let us suppose that a lens is placed before the aperture, as in Fig. 113, so that all the light which leaves the aperture parallel to a given direction is focussed at a single point.

[2] The angles are taken this way in order to embrace the case in which the plane of incidence is not the same as the plane of diffraction. Here $i$ and $\theta$ are the angles which the rays make with a plane perpendicular to AB.

intensity of the diffracted light consequently depends on the final position of the tracing point M—that is, upon the phase difference of the extreme rays AX and BY. For example, when this phase difference is an even multiple of $\pi$ the intensity is zero, for then the tracing point coincides with O, and OM is zero. Now if we denote the phase difference[1] of the extreme rays by $2\phi$ we have $2\phi = \text{OCM} = 2\pi\delta/\lambda$, and therefore

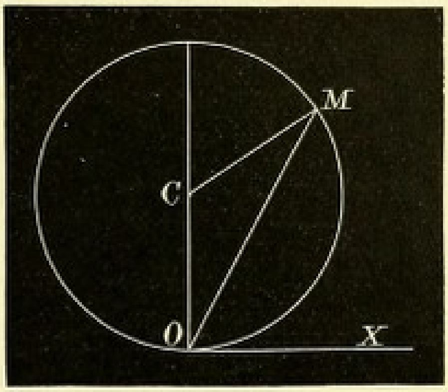



Fig. 132.

$$\phi = \frac{\pi a}{\lambda}(\sin i + \sin \theta).$$

But we also have

$$\text{OM} = 2\text{R} \sin \tfrac{1}{2}\text{OCM} = 2\text{R} \sin \phi,$$

where R is the radius of the circle. Further, if $s$ be the length of the arc OM, then by the manner in which the curve is plotted $s$ must be proportional to the width of the aperture.

So that if we take the constant of proportionality to be unity we may write

$$s = a.$$

But $s = 2\text{R}\phi$, therefore $2\text{R} = a/\phi$, and the expression for OM becomes

$$\text{OM} = a\frac{\sin \phi}{\phi},$$

and the resultant intensity is measured by

$$\text{I} = a^2\frac{\sin^2\phi}{\phi^2}.$$

*Cor.* 1.—Since the line OM makes with OX an angle $\text{MOX} = \frac{1}{2}\text{MCO}$, it follows that the phase of the resultant vibration is the same as that of the vibration contributed by the middle strip of the aperture. Hence if the vibration from B be represented by $y = \sin \omega t$, that from A will be $y = \sin(\omega t + 2\phi)$, and the equation of the resultant vibration is

$$y = a\frac{\sin \phi}{\phi}\sin(\omega t + \phi).$$

[1] It is to be remembered that the tracing point may have described the circle several times, and finally settled in the position M. The angle $2\phi$ is the whole angle through which the radius CM has revolved, and the arc $s$ referred to in the text is the whole length of arc described by M.

It is also worthy of note that R varies from point to point of the screen. It is $s$ or the width of the aperture that remains constant, and it is for this reason that the maxima are not determined by $\phi$ = an odd multiple of $\pi/2$, as shown in Art. 151.

*Cor.* 2.—If the light be incident perpendicularly on the aperture, the intensity at any point of the diffraction pattern is proportional to

$$a^2 \frac{\sin^2\left(\frac{\pi a}{\lambda}\sin\theta\right)}{\left(\frac{\pi a}{\lambda}\sin\theta\right)^2}.$$

**151. Determination of the Maxima and Minima.**—The expression for the intensity at any point, being a function of the angle $\theta$ of diffraction, will vary from point to point on the screen, and pass through a series of maximum and minimum values. This has already been indicated by the elementary examination of Art. 129, but we are now in a position to inquire into the phenomena more accurately.

For brevity we have written $\phi = \frac{\pi a}{\lambda}(\sin i + \sin\theta)$, and we have found the intensity measured by $a^2 \sin^2\phi/\phi^2$, hence as the angle $\theta$ varies, the intensity passes through a series of maximum and minimum values as follows.

*Minima.*—$\mathrm{I} = 0$ when $\phi = n\pi$, excluding the value $n = 0$, which corresponds to a maximum. Hence at points in the direction $\theta$ determined by the equation

$$\sin i + \sin\theta = n\lambda/a$$

there is complete darkness when $n$ is any integer other than zero.

*Maxima.*—Equating to zero the first derived of $\sin\phi/\phi$, we find that the values of $\phi$ which make I a maximum satisfy the equation

$$\phi = \tan\phi.$$

To solve this equation graphically, plot first the curves

$$(1)\ y = x, \quad \text{and } (2)\ y = \tan x.$$

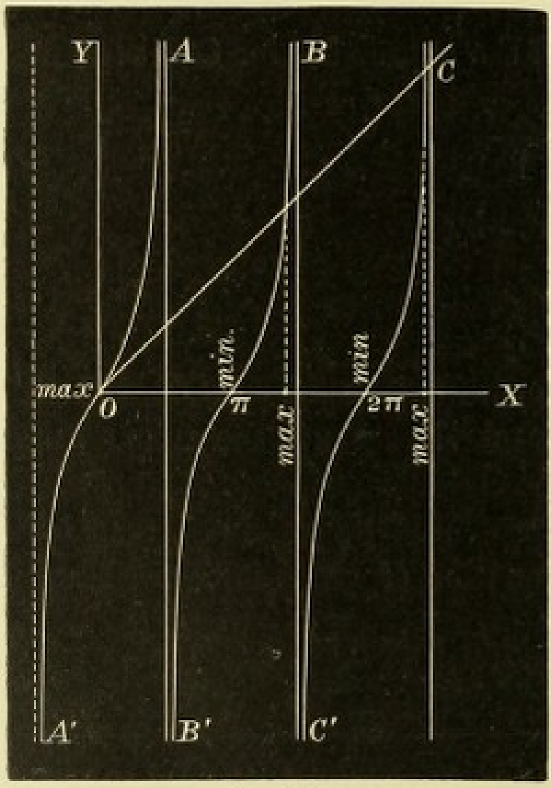


Fig. 133.

The first represents a line bisecting the angle between the axes of $x$ and $y$ (Fig. 133). The second consists of an infinite number of branches of amplitude $\pi$. The first branch AA′ passes through the origin O and touches at infinity the lines $x = \pm\frac{1}{2}\pi$, which are its asymptotes. The second branch BB′ cuts the axis of $x$ at the point $x = \pi$ and touches at infinity the lines $x = \frac{3}{2}\pi$ and $x = \frac{1}{2}\pi$. The maximum values of I correspond to the values of $y$ which satisfy both these equations, since then we have $x = \tan x$. The maxima are therefore determined by the intersection of the line OC with the curves AA′, BB′, CC′, etc. The corresponding values of $\phi$ are less than the odd multiples of $\frac{1}{2}\pi$, but as $n$ increases the value of $\phi$ approaches more and more nearly to $(2n+1)\frac{1}{2}\pi$. The values of $\phi$ corresponding to the maxima values of the illumination have been given by Schwerd as follows :—

| | |
|---|---|
| $\phi_0 = 0$ | $I_0 = 1$ |
| $\phi_1 = 1 \cdot 4303\pi$ | $I_1 = \sin^2\phi_1/\phi_1^2$ |
| $\phi_2 = 2 \cdot 4590\pi$ | $I_2 = \sin^2\phi_2/\phi_2^2$ |
| $\phi_3 = 3 \cdot 4709\pi$ | $I_3 = \sin^2\phi_3/\phi_3^2$ |
| $\phi_4 = 4 \cdot 4774\pi$ | $I_4 = \sin^2\phi_4/\phi_4^2$ |
| $\phi_5 = 5 \cdot 4818\pi$ | $I_5 = \sin^2\phi_5/\phi_5^2$ |
| $\phi_6 = 6 \cdot 4844\pi$ | $I_6 = \sin^2\phi_6/\phi_6^2$ |
| $\phi_7 = 7 \cdot 4865\pi$ | $I_7 = \sin^2\phi_7/\phi_7^2$ |

Thus while $\phi_n$ becomes more nearly $(2n+1)\frac{1}{2}\pi$ as $n$ increases the corresponding values of I (the maxima illuminations) decrease rapidly, being approximately in the ratios of the quantities 1, $\left(\frac{2}{3\pi}\right)^2$, $\left(\frac{2}{5\pi}\right)^2$, etc., which correspond to $\phi$, equal to odd multiples of $\pi$. If the intensity of the first maximum be taken as unity, the values of the second, third, and fourth will be approximately $\frac{1}{20}$, $\frac{1}{50}$, and $\frac{1}{110}$ respectively. Fig. 134 represents the variations of intensity, its abscissæ being the angle $\phi$ and its ordinates the corresponding intensities. The first maximum is very much greater than the others, and these again diminish very rapidly. With white light we have a series of rainbow-coloured fringes, violet at their inner and red at their outer edges. The spectra formed by a single narrow aperture Fraunhofer terms *spectra of the first class.*

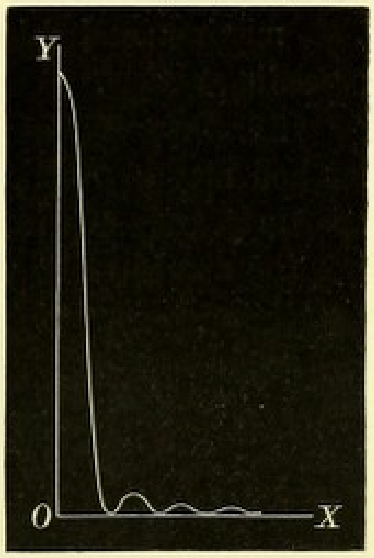


Fig. 134.

**152. Circular Aperture.**—The intensity at any point on the axis of a circular aperture may be easily expressed in the same manner. Thus we have already seen (Art. 53) that when the aperture is divided into circular annuli, corresponding to equal differences of phase, the amplitudes of the vibrations produced by these elements are equal, and therefore (when the influence of obliquity is neglected) the vibration curve is a circle. Hence, as in Art. 150, if $s$ be the length of the arc of the circle which represents the resultant vibration, we have

$$OM = 2R \sin \phi, \quad \text{and } s = 2R\phi.$$

Therefore the intensity may be expressed in the form

$$I = s^2 \frac{\sin^2\phi}{\phi^2},$$

where $2\phi$ is the phase difference of the vibrations from the centre and the circumference of the aperture.

When the radius of the aperture is small compared with the other distances involved we have approximately, as in Art 130,

$$2\phi = \frac{2\pi\delta}{\lambda} = \frac{2\pi}{\lambda}\left(a + \frac{r^2}{2a} + b + \frac{r^2}{2b} - a - b\right) = \frac{\pi r^2(a+b)}{ab\lambda},$$

where $r$ is the radius of the aperture, $a$ its distance from the luminous origin, and $b$ its distance from the screen. Now by Art. 53 the arc $s$ is proportional to $2\pi a\delta/(a+b)$, and this by the foregoing reduces to $\pi r^2/b$. Hence the expression for I becomes

$$I \propto \left(\frac{\pi r^2}{b}\right)^2 \frac{\sin^2\phi}{\phi^2},$$

where

$$\phi = \pi r^2(a+b)/2ab\lambda.$$

The maxima and minima in this case may be discussed as in the preceding article.

**153. Two Equal Rectilinear Apertures.**—The same graphic method may be applied with facility to the case of two very narrow apertures, each of width $a$ separated by an opaque interval of width $b$. Let the apertures be OA and BC (Fig. 135). Then, as in Art. 150, the effect of each aperture may be represented by an arc of a circle of magnitude $2\alpha$, where

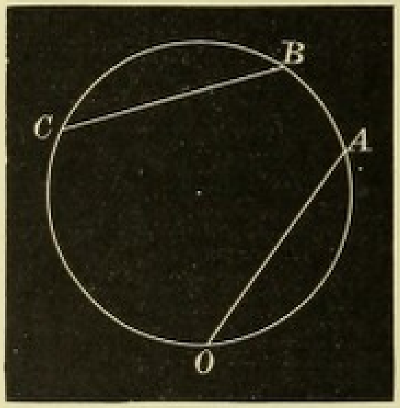


Fig. 135.

$$\alpha = \frac{\pi a}{\lambda}(\sin i + \sin \theta),$$

and these arcs (Fig. 135) will be separated by an arc AB of magnitude $2\beta$ given by the equation

$$\beta = \frac{\pi b}{\lambda}(\sin i + \sin \theta).$$

The resultant amplitude due to each will be measured by the chord OA or BC—that is, by $a \sin \alpha/\alpha$, but as the resultant vibrations due to OA and BC differ in phase by an amount $2(\alpha + \beta)$, namely, the angular distance between their middle points, it follows by Art. 43 that the resultant intensity is measured by

$$I = (2a)^2 \frac{\sin^2 \alpha}{\alpha^2} \cos^2(\alpha + \beta).$$

The equation of the resultant vibration is therefore

$$y = 2a\frac{\sin\alpha}{\alpha}\cos(\alpha+\beta)\sin(\omega t + 2\alpha + \beta),$$

and this corresponds in phase to the middle strip of the opaque interval. For the resultant vibration transmitted by OA is

$$y_1 = a\frac{\sin\alpha}{\alpha}\sin(\omega t + \alpha)$$

and that transmitted by BC is

$$y_2 = a\frac{\sin\alpha}{\alpha}\sin(\omega t + 3\alpha + 2\beta).$$

But $y = y_1 + y_2$, therefore, etc.

The intensity depends on two variable factors, one $\sin\alpha/\alpha$, which gives the fringes of a single aperture, and the other $\cos(\alpha+\beta)$, which gives a system of fringes corresponding to the interference of the lights from the two apertures. This factor vanishes when

$$(\alpha+\beta) = (2n+1)\frac{\pi}{2},$$

that is, when

$$(a+b)(\sin i + \sin\theta) = (2n+1)\frac{\lambda}{2}.$$

In this case the light from the second aperture is an odd number of half-period elements behind that from the first, and the two destroy each other by interference. But if

$$\alpha + \beta = n\pi,$$

or

$$(a+b)(\sin i + \sin\theta) = n\lambda,$$

then the two are concordant, and the illumination is a maximum. These are termed maxima and minima of the second order, or *spectra of the second class.*

We may therefore consider the phenomena observed as the superposition of these two systems of fringes; that due to the first factor, a diffraction system, and that due to the second factor, an interference system. The intensity is zero when either factor vanishes.

The dispersion of the second system being inversely as $a + b$ is less than that of the first system, which is inversely as $a$, it follows that when the apertures are not very close the second system is nearly all contained within the first two bands of the first system.

The points of maximum intensity consequently do not in general coincide either with those of the first system or with those of the second.

**154. Small Rectangular Aperture.**—In the calculation of Art. 150 the diffraction pattern is considered only in so far as it depends on the width of the aperture. In the case of a long narrow slit the pattern consists of a system of rectilinear bands parallel to the length of the slit, and these bands arise from the narrowness of the aperture. The length of the slit, in fact, is so great that all diffraction effects in this dimension are lost, for the whole effective portion of a strip of the wave taken parallel to the length of the slit is transmitted when the aperture is long.

On the other hand, when the aperture is short as well as narrow, so as to have the shape of a small rectangle of length $a$ and width $b$, then the limited length comes into operation and produces diffraction effects. The whole effective portion of a wave strip taken parallel to the length is not transmitted, but partly obstructed, by the aperture. The pattern consists in fact of a system of bands parallel to the length of the aperture, and also a system of bands parallel to the width. The former arise from the limited width of the aperture, and the latter from its limited length. In order to determine the intensity let the aperture be divided into a great number of very small strips parallel to its length. Each of these strips will be of length $a$, and will give rise to a vibration at any point under consideration of amplitude (Art. 150)

$$\mathrm{A} = a\frac{\sin \alpha}{\alpha}, \quad \text{where } \alpha = \frac{\pi a}{\lambda}(\sin i + \sin \theta),$$

and the incident light makes an angle $90^\circ - i$ with the length of the slit, while the diffracted light makes an angle $90^\circ - \theta$. We have now to find the result of a great number of vibrations of amplitudes $\mathrm{A} = a \sin \alpha/\alpha$ and varying in phase from $\omega t$ to $\omega t + 2\beta$, where

$$\beta = \frac{\pi b}{\lambda}(\sin i' + \sin \theta'),$$

$90^\circ - i'$ and $90^\circ - \theta'$ being the angles which the incident and diffracted light make with the direction of the width of the slit. But as before, the resultant amplitude of these will be $\mathrm{A}b \sin \beta/\beta$. Hence the intensity of illumination will be measured by

$$\mathrm{I} = a^2 b^2 \frac{\sin^2 \alpha}{\alpha^2} \frac{\sin^2 \beta}{\beta^2}.$$

The illumination at any point therefore depends on two variable factors, one of which gives rise to a series of bands parallel to the side $a$ of the aperture, while the other gives a series of bands parallel to the side $b$ of the aperture. These lines enclose a system of rectangles (Fig. 136) similar to the aperture turned through 90°. The greater the length of a side the narrower are the bands perpendicular to that side. It is thus we have only one system of parallel fringes with a long narrow slit, for the width of the slit is so small that the bands parallel to the length are fairly broad, while those parallel to the width are invisible on account of the length of the slit.

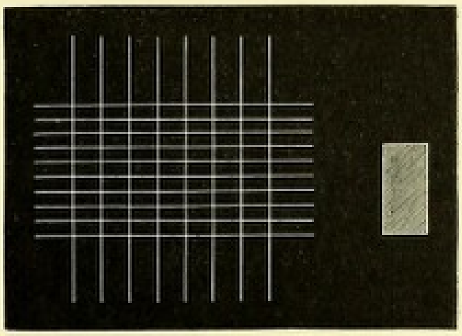


Fig. 136.

**155. Talbot's Bands.**—The system of bands which are seen crossing a tolerably pure spectrum, when it is viewed through a small hole half covered with a thin transparent plate, has been mentioned already in Art. 106. We are now in a position to easily deduce the expression for the illumination at any point, the aperture being supposed rectangular.

Let AB (Fig. 137) represent the plate covering half the aperture AC. The effect of the illumination from AB will be represented by an arc OA = $2\alpha$ (Fig. 138) of a circle, but as the plate produces a retardation, it follows that the ray from B (Fig. 137) which does not

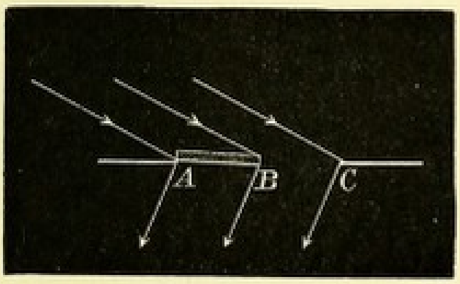



Fig. 137.

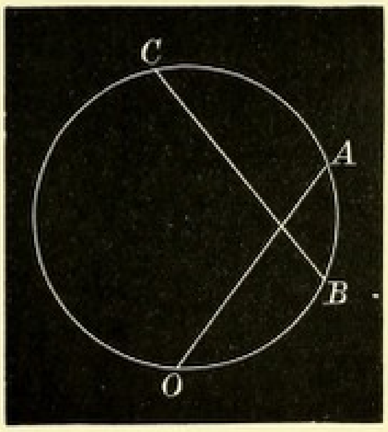



Fig. 138.

traverse the plate will be accelerated relatively to that which passes through the plate by some amount $2\delta$. Hence the effect of the free part BC of the aperture will be represented by the arc BC = $2\alpha$ (Fig. 138) of the circle, where AB = $2\delta$, the retardation in the plate. It follows easily that the inclination of the chords OA and BC is $2(\alpha - \delta)$, consequently if they are each equal to $\rho$, their resultant is

$$2\rho \cos (\alpha - \delta).$$

But by the preceding article $\rho = ab \sin\alpha \sin\beta/\alpha\beta$, where $ab$ is the area of the aperture, consequently the resultant illumination is proportional to

$$\frac{\sin^2\alpha}{\alpha^2}\frac{\sin^2\beta}{\beta^2}\cos^2(\alpha-\delta),$$

which is the formula deduced analytically by Airy.[1] A table of the values of this expression for various values of $\alpha$ and $\beta$ was constructed by Airy, and he also plotted curves showing the fluctuations of intensity.

**156. The Diffraction Grating—Any Number of Parallel, Equal, and Equidistant Narrow Rectangular Apertures.**—In the case of a system of $n$ very narrow equal apertures, separated by equal opaque intervals of width $b$, we have to find the resultant of a system of amplitudes represented by the chords of $n$ arcs of a circle each of magnitude $2\alpha$, and separated from each other by arcs each equal to $2\beta$ (Fig. 139).

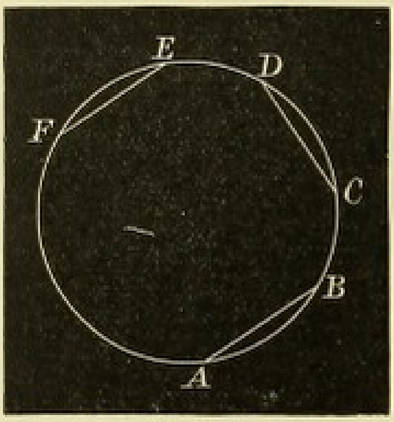


Fig. 139.

If we take any axes of reference OX and OY, and if any one of the chords makes an angle $\eta$ with OX, the consecutive chord will make an angle $\eta + 2\alpha + 2\beta = \eta + \gamma$ suppose, and the other chords will make angles $\eta + 2\gamma$, $\eta + 3\gamma$, . . . $\eta + (n-1)\gamma$ respectively with OX. Hence if X denotes the sum of the projections of all the chords on the axis OX, and if $\rho$ be the length of each chord, we have

$$\begin{aligned}X &= \rho[\cos\eta + \cos(\eta+\gamma) + \cos(\eta+2\gamma) + \ . \ . \ . \ \cos\{\eta+(n-1)\gamma\}],\\ &= \rho\frac{\cos\{\eta+\frac{1}{2}(n-1)\gamma\}\sin\frac{1}{2}n\gamma}{\sin\frac{1}{2}\gamma}.\end{aligned}$$

In the case of a long narrow aperture (Art. 150), such as we are now considering, $\rho$ is given by the equation

$$\rho = a\frac{\sin\alpha}{\alpha}.$$

Similarly if Y denotes the sum of the projections on OY, we have

$$\begin{aligned}Y &= \rho[\sin\eta + \sin(\eta+\gamma) + \sin(\eta+2\gamma) + \ . \ . \ . \ \sin\{\eta+(n-1)\gamma\}],\\ &= \rho\frac{\sin\{\eta+\frac{1}{2}(n-1)\gamma\}\sin\frac{1}{2}n\gamma}{\sin\frac{1}{2}\gamma}.\end{aligned}$$

[1] Airy, *Phil. Trans.* p. 1, 1841.

Hence [1]

$$X^2 + Y^2 = \rho^2 \frac{\sin^2 \frac{1}{2}n\gamma}{\sin^2 \frac{1}{2}\gamma}.$$

But $X^2 + Y^2$ is the square of the resultant amplitude. Consequently the resultant intensity is measured by

$$I = \rho^2 \frac{\sin^2 n(\alpha+\beta)}{\sin^2 (\alpha+\beta)},$$

where $\rho^2$ is the intensity produced by a single aperture (Arts. 152-154).

*Cor.*—The phase $\phi$ of the resultant vibration is given by the equation

$$\tan \phi = \frac{Y}{X} = \tan\{\eta + \tfrac{1}{2}(n-1)\gamma\},$$

$$= \tan\{\eta + (n-1)(\alpha+\beta)\},$$

it is consequently the same as the phase of the vibration from the middle point of the grating. Hence if the equation of the vibration from the first aperture of the grating be $y = \rho \sin \omega t$, the equation of the resultant vibration will be

$$y = \rho \frac{\sin n(\alpha+\beta)}{\sin(\alpha+\beta)} \sin\{\omega t + (n-1)(\alpha+\beta)\}.$$

**157. Determination of the Maxima and Minima Intensities.**—The expression for the intensity of the illumination produced at any point by a grating is the product of two variable factors; one $\rho^2$

[1] Otherwise thus denoting $\sqrt{-1}$ by $i$, we have, since $\cos\theta + i\sin\theta = e^{i\theta}$,

$$X + iY = \rho[e^{i\eta} + e^{i(\eta+\gamma)} + e^{i(\eta+2\gamma)} + \ldots] = \frac{\rho e^{i\eta}(1 - e^{in\gamma})}{1 - e^{i\gamma}}.$$

Similarly

$$X - iY = \rho[e^{-i\eta} + e^{-i(\eta+\gamma)} + \ldots] = \frac{\rho e^{-i\eta}(1 - e^{-in\gamma})}{1 - e^{-i\gamma}}.$$

Multiplying we find

$$X^2 + Y^2 = \rho^2 \frac{(2 - e^{in\gamma} - e^{-in\gamma})}{(2 - e^{i\gamma} - e^{-i\gamma})}, \quad = \rho^2 \frac{(1 - \cos n\gamma)}{1 - \cos\gamma} = \rho^2 \frac{\sin^2 \frac{1}{2}n\gamma}{\sin^2 \frac{1}{2}\gamma}.$$

So also by addition we find

$$X = \rho \frac{\cos\{\eta + \frac{1}{2}(n-1)\gamma\} \sin \frac{1}{2}n\gamma}{\sin \frac{1}{2}\gamma},$$

and by subtraction

$$Y = \rho \frac{\sin\{\eta + \frac{1}{2}(n-1)\gamma\} \sin \frac{1}{2}n\gamma}{\sin \frac{1}{2}\gamma}.$$

corresponds to the diffraction produced by a single aperture, and has been already discussed. The other factor,

$$\frac{\sin^2 n(\alpha+\beta)}{\sin^2 (\alpha+\beta)},$$

also produces a series of maxima and minima corresponding to the interference of the light from the various apertures. For brevity let us write $x = \alpha + \beta$, then the bright and dark bands are determined by equating to zero the first derived of $\sin^2 nx/\sin^2 x$—that is, by

$$\frac{2 \sin nx}{\sin^3 x}(n \sin x \cos nx - \cos x \sin nx) = 0,$$

which is satisfied by

$$(1)\ \sin nx = 0, \quad (2)\ n \tan x = \tan nx.$$

*Minima.*—In the first case if $\sin\ nx = 0$, we have $nx = m\pi$, and $\sin\ n(\alpha + \beta)/\sin\ (\alpha + \beta) = 0$, so that the amplitude vanishes and we have a series of minima of zero value.

*Principal Maxima.*—If, however, $x = m\pi$, both the numerator and denominator of the expression $\sin nx/\sin x$ will vanish. Its true value, however, will be $n$, so that the intensity will be a maximum and proportional to $n^2$. This corresponds to

$$\alpha + \beta = m\pi, \quad \text{or } (a+b)(\sin i + \sin \theta) = m\lambda.$$

These maxima are very intense and are termed principal maxima. There are obviously $n - 1$ minima between two principal maxima.

*Secondary Maxima.*—The roots of the equation

$$n \tan x = \tan nx$$

other than $x = m\pi$ (which correspond to the principal maxima) give rise to another set of maxima termed secondary maxima, much less intense than the principal maxima. From the equation $n$ tan $x =$ tan $nx$ we find

$$\frac{\sin^2 nx}{\sin^2 x} = \frac{n^2}{1 + (n^2 - 1)\sin^2 x},$$

which shows that the ratio of these secondary maxima to the principal maxima ($n^2$) is

$$\frac{1}{1 + (n^2 - 1)\sin^2 x},$$

and when $n$ is large these secondary fringes are very weak and entirely lost, but when $n$ is small they may be observed. Fig. 140

shows the principal maxima with four secondary maxima between each pair. Since we have $n-1$ minima between two principal maxima, it follows that we have $n-2$ secondary maxima between each pair of principal maxima. The figure is constructed to represent the case $n=6$. The secondary maxima are unequal, not equidistant, and small compared with the principal maxima, especially if the number of apertures is large.

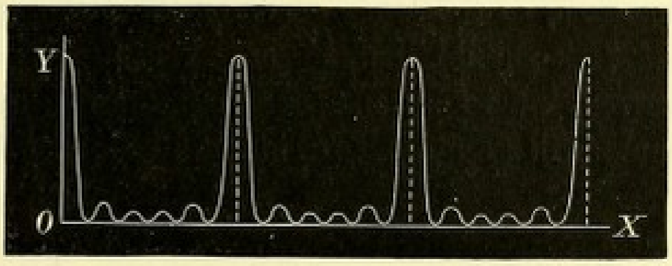


Fig. 140.

Hence if the illumination depended only on the factor $\sin^2 nx/\sin^2 x$ we should be presented with a set of bright bands of equal intensity proportional to $n^2$ where $n$ is the number of lines in the grating. These are the principal maxima. Between each pair of them we have a set of narrow fringes which become more and more narrow and indistinct as the number of apertures is increased. Consequently with a large number of apertures, as in the diffraction grating, they are not discernible.

The secondary maxima may be determined by the intersections of the curves

$$(1)\ y = n \tan x, \quad \text{and } (2)\ y = \tan nx,$$

in a manner analogous to that employed in the case of a single aperture (Art. 151).

The first equation represents a curve asymptotic to the line $x=\frac{1}{2}\pi$ while the second is a similar curve, or a set of similar curves asymptotic to $nx=\frac{1}{2}\pi$. Fig. 141 represents the case of $n=6$. The symmetry of $x$ shows that the effect of this factor remains unaltered if the opaque portions be made transparent and the transparent portions opaque, as this merely amounts to interchanging $a$ and $\beta$.

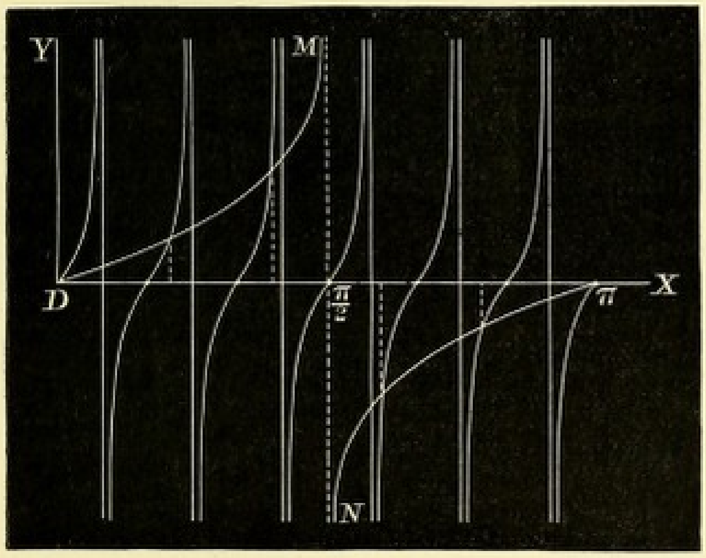


Fig. 141.

The resultant illumination at any point being determined by the product of the two factors $\sin^2 a/a^2$ and $\sin^2 nx/\sin^2 x$, to obtain it we must multiply the ordinates of the curve (Fig. 140) by the corresponding ordinates of the curve relative to a single aperture (Fig. 134). The variations of the latter are very feeble compared with the principal maxima, so that they scarcely affect the appearance.

If it should happen, however, that a zero value of $\sin \alpha/\alpha$ should correspond to a principal maximum of the other curve, then this maximum will be absent. This will happen when

$$a = m\lambda, \quad \text{and also } (a+b) = m'\lambda,$$

or

$$\frac{a}{b} = \frac{m}{m' - m},$$

where $m$ and $m'$ are whole numbers, and the corresponding principal maxima, or spectra, are absent (see Art. 140).

**158. Any Number of Narrow Rectangular Apertures, Parallel and Equal, but not Equidistant.**—The foregoing calculation is based on the supposition that the opaque intervals are of equal width, or, in other words, that the grating has been ruled uniformly; the investigation may, however, be applied with the greatest facility to the case in which the length $b$ is variable, which is that of a system of equal apertures placed at random distances apart.

Each aperture will be represented by an arc $2\alpha$ of a circle, and these arcs will be arranged at random round the circumference of the circle—that is, separated by variable intervals. Let the chords of the arcs $2\alpha$ make angles $\eta_1, \eta_2, \ldots \eta_n$ with a fixed axis OX, and let each chord be $\rho$ as before, then the sum of the projections of the chords on OX is

$$X = \rho(\cos \eta_1 + \cos \eta_2 + \cos \eta_3 + \ldots \cos \eta_n),$$

and the sum of their projections on the perpendicular axis OY is

$$Y = \rho(\sin \eta_1 + \sin \eta_2 + \sin \eta_3 + \ldots \sin \eta_n).$$

Hence the square of the resultant amplitude is

$$X^2 + Y^2 = \rho^2[n + 2\Sigma \cos (\eta_1 - \eta_2)].$$

Now the sum $\Sigma \cos (\eta_1 - \eta_2)$ embraces all combinations of the angles $\eta_1, \eta_2, \ldots \eta_n$, and the values of its constituents vary irregularly, taking random values between $+1$ and $-1$. Admitting that their sum is negligible or sensibly zero (see Example, p. 48), we have

$$X^2 + Y^2 = n\rho^2,$$

or the intensity is $n$ times that produced by a single aperture. Substituting for $\rho$ from Art. 154, we have

$$I = na^2b^2 \frac{\sin^2 \alpha}{\alpha^2} \cdot \frac{\sin^2 \beta}{\beta^2}.$$

**159. General Investigation.**—The components of the resultant vibration at any point and the intensity of the illumination may be very easily expressed by means of the vibration spiral. Thus if $x$ and $y$ be the co-ordinates of any point on the spiral, and $\phi$ the inclination of the tangent at that point to the axis of $x$, then $\phi$ is the phase of the vibration from the corresponding point of the wave.

Now in any curve we have, if $ds$ be the element of arc,

$$dx = \cos\phi ds, \quad \text{and } dy = \sin\phi ds.$$

Therefore

$$x = \int \cos\phi ds, \quad \text{and } y = \int \sin\phi ds,$$

and consequently the radius vector is given by the equation

$$r^2 = x^2 + y^2 = [\int \cos\phi ds]^2 + [\int \sin\phi ds]^2.$$

But the resultant amplitude is given by the radius vector $r$ of the spiral, consequently the intensity of illumination is measured by

$$\mathrm{I} = [\int \cos\phi ds]^2 + [\int \sin\phi ds]^2.$$

The phase $\phi_0$ of the resultant vibration is given by the inclination of $r$ to the axis of $x$, and we have consequently

$$\tan\phi_0 = \frac{y}{x} = \frac{\int \sin\phi ds}{\int \cos\phi ds},$$

the integrals being taken between limits which are determined by the problem under investigation.

Now the element of arc $ds$ of the spiral is taken proportional to the amplitude of the vibration contributed by an element of area $d\mathrm{S}$ of the wave front, and this amplitude is taken proportional to the area and inversely as the distance ($r$) of the element from the point under consideration. Further, the amplitude will depend on the inclination of $r$ to the wave normal—that is, on the obliquity, as well as on the direction of the vibration in the wave front. The full expression for $ds$ is consequently of the form

$$ds = \frac{d\mathrm{S}}{r} \cdot f.,$$

where $f$ is some function of the co-ordinates, which diminishes as the element recedes from the pole of the wave.

The general expressions for the resultant consequently take the form

$$x = \int f \cos\phi \frac{d\mathrm{S}}{r}, \quad y = \int f \sin\phi \frac{d\mathrm{S}}{r}.$$

The intensity of the illumination is consequently

$$I=\left[\int f\cos\phi\frac{dS}{r}\right]^2+\left[\int f\sin\phi\frac{dS}{r}\right]^2,$$

while the phase of the resultant is

$$\tan\phi_0=\frac{y}{x}=\frac{\int f\sin\phi\frac{dS}{r}}{\int f\cos\phi\frac{dS}{r}}.$$

In the foregoing investigation the amplitude of the vibration contributed by an element $dS$ of the wave front, situated at a distance $r$ has been represented by the expression $fdS/r$, and it is consequently desirable to investigate the constitution of the symbol $f$. Now if A be taken to represent the amplitude of the vibration in the wave front, then the amplitude produced by an element $dS$ at a distance $r$ may be written in the form

$$a=\frac{AdS}{r}b\psi,$$

where $\psi$ is some function of the angles which $r$ makes with the wave normal and the direction of vibration, and $b$ is a quantity which has yet to be determined. Now $a$ and A are of the same dimensions, being both amplitudes, whereas $\psi$ is a function of angles and is of zero dimensions, consequently $b$ must be the inverse of a length. But the only other quantities that can have any reference to $a$ are $v$ and $\lambda$, and of these $v$ is eliminated by the fact that it is a function of the time; hence the only length that can enter into the quantity $b$ is the wave length $\lambda$. We conclude, therefore, that $b$ is of the form $k/\lambda$ where $k$ is a constant of zero dimensions. Using this notation the complete expression for the amplitude is of the form

$$a=\frac{kAdS}{r\lambda}\psi.$$

Comparing this with the form $fdS/r$, we find that $f$ is of the form

$$f=\frac{kA}{\lambda}\psi$$

where $k$ may be taken as unity. We have thus reached the important result that the factor $f$ contains the reciprocal of the wave length as a constituent as well as depending on the obliquity and the amplitude of vibration in the original wave front.

## *Examples*

1. If the radius vector $r$, from an external point O, to any element of a plane wave makes an angle $\theta$ with the wave normal, the components of the vibration at O are

$$x = k\int \cos(2\pi bu/\lambda) \,.\, f \,.\, du,$$

$$y = k\int \sin(2\pi bu/\lambda) \,.\, f \,.\, du,$$

where $u = \sec\theta - 1$, and $k$ is a constant.

[Divide the wave up into circular elements around the pole P as centre (Fig. 24). Let $\rho$ be the radius of one of these elements and $r$ its distance from O. Then if $OP = b$, we have $\rho = b\tan\theta$ and $r = b\sec\theta = b(1+u)$,

$$\phi = \frac{2\pi}{\lambda}(r - b) = \frac{2\pi b}{\lambda}(\sec\theta - 1) = \frac{2\pi bu}{\lambda},$$

$$dS = 2\pi\rho d\rho = 2\pi b^2 \tan\theta \sec^2\theta d\theta = 2\pi b^2(1+u)du.$$

Therefore $dS/r = 2\pi b du$, etc.]

2. If $f$ = constant, show that the vibration curve is a circle.

[Here we have

$$x = k\int_0^u \cos(2\pi bu/\lambda)du = a\sin(2\pi bu/\lambda),$$

$$y = k\int_0^u \sin(2\pi bu/\lambda)du = a\{1 - \cos(2\pi bu/\lambda)\}$$

where $a$ is a constant.

Hence

$$x^2 + (y-a)^2 = a^2,$$

which shows that the amplitude curve is a circle, of radius $a$, passing through the origin and having its centre situated on the axis OY (cf. Art. 150).]

3. In the same case prove that the vibration excited at O by the complete wave is in phase a *quarter period* behind that which reaches it from the pole P.

[For the complete wave we have, if we assume that $\cos\infty = \sin\infty = 0$ (?)

$$x = k\int_0^\infty \cos(2\pi bu/\lambda)du = 0,$$

$$y = k\int_0^\infty \sin(2\pi bu/\lambda)du = k.$$

Hence if $\phi$ be the phase of the resultant vibration, we have

$$\tan\phi = y/x = \infty,$$

therefore

$$\phi = \tfrac{1}{2}\pi.]$$

4. If the effective portion of a wave be confined to a small portion around the pole, which is sensibly plane, and for which $f$ is constant, the amplitude curve will be a circle.

[We have

$$dS = 2\pi\rho d\rho = \pi d\rho^2,$$

and

$$\delta=\frac{2\pi}{\lambda}\cdot\frac{\rho^2}{2b}, \quad \text{since } r=b+\frac{\rho^2}{2b}.$$

Hence

$$dS=b\lambda d\delta,$$

and therefore, since $f$ is supposed constant,

$$x=a\int_o^\delta \cos\delta d\delta=a\sin\delta,$$

$$y=a\int_o^\delta \sin\delta d\delta=a(1-\cos\delta),$$

and

$$x^2+(y-a)^2=a^2, \text{ etc.]}$$

5. Determine the radius of curvature of the amplitude curve, Ex. 1.

[The amplitude of the vibration excited by the element $dS$ is proportional to

$$\frac{f\cdot dS}{r}=2\pi bf\cdot du.$$

Hence the element $ds$ of the amplitude curve is proportional to $fdu$, and if $\phi$ be the angle the tangent to it makes with OX, we have $d\phi=2\pi bdu/\lambda$. Hence the radius of curvature $\rho$ is proportional to

$$\rho=\frac{ds}{d\phi}=\lambda\cdot f.$$

Hence if $f$=constant the amplitude curve is a circle, and if $f$ diminishes gradually from the pole of the wave, the curve is a spiral of ever decreasing radius of curvature as we move along it, setting out from O.]

6. If

$$f=1+\cos\theta=\frac{2+u}{1+u},$$

we have for a plane wave (**Ex.** 1)

$$x=k\int\cos(2\pi bu/\lambda)\frac{2+u}{1+u}du,$$

$$y=k\int\sin(2\pi bu/\lambda)\frac{2+u}{1+u}du.$$

**160. Second Graphic Method.**—A second method of representing the resultant amplitude of a system of superposed vibrations has been given in Art. 45. This method is based on a second graphic representation of the resultant of a system of forces, and it follows that if a system of lines $OP_1$, $OP_2$, etc. (Fig. 12), be drawn from any point O such that the lengths $OP_1$, $OP_2$, etc., represent the amplitudes of the vibrations, and the angles they make with a fixed line OX their phases, the resultant will be represented by $n$ times the line OG, joining O to the centre of mean position (G) of the points $P_1$, $P_2$, etc.

If in this manner we plot down the curve for a complete wave, beginning at the pole of the wave, we see at once that we have a spiral curve surrounding O with many convolutions of ever decreasing radius (Fig. 142). With this spiral the resultant effect of any portion of the wave is not represented by the chord joining O to a point on the curve, but by that joining it to the centre of gravity of the corresponding arc of the curve. This obviously passes through fluctuations as in the case of the other spiral.

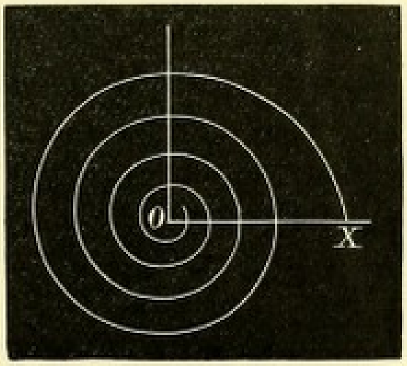


Fig. 142.

This method may be applied with the greatest facility to the calculation of the intensity of the illumination at any point of a diffraction pattern when the incident light is parallel. In this case the curve becomes a circle, and the problem is reduced at once to the finding of the centre of gravity of an arc of a circle, leading to all the results we have already arrived at. The deduction of these results by this method will form a simple and useful exercise.

The equation of the second spiral and the integrals giving the general expression for the intensity may be derived very simply. For if $\bar{x}$ and $\bar{y}$ be the co-ordinates of the centre of mean position of the extremities of the radii vectores of the spiral, we have

$$n\bar{x} = \Sigma x \quad \text{and} \quad n\bar{y} = \Sigma y.$$

Hence

$$(n \,.\, \text{OG})^2 = (n\bar{x})^2 + (n\bar{y})^2 = (\Sigma x)^2 + (\Sigma y)^2.$$

Now any radius vector $\rho$ of the spiral represents one of the constituent vibrations in amplitude and phase. The length of $\rho$ is consequently (as in Art. 159) given by the equation

$$\rho = \frac{dS}{r} \cdot f.$$

But if $x$ and $y$ be the co-ordinates of the extremity of $\rho$ we have

$$x = \rho \cos \phi, \quad y = \rho \sin \phi,$$

where $\phi$ is the phase of the corresponding vibration. Hence the expression for the intensity of the illumination

$$(n \,.\, \text{OG})^2 = (\Sigma x)^2 + (\Sigma y)^2 = (\Sigma \rho \cos \phi)^2 + (\Sigma \rho \sin \phi)^2$$

gives as before

$$\text{I} = \left[\int f \cos \phi \frac{dS}{r}\right]^2 + \left[\int f \sin \phi \frac{dS}{r}\right]^2.$$

**161. Fresnel's Integrals.**—In the foregoing articles it has been proved that the intensity of the illumination at any point of a diffraction pattern may be expressed in general as the sum of the squares of two integrals taken between limits defined by the nature of the problem. By making certain assumptions and approximations Fresnel deduced special forms of these integrals which we shall now consider. Let us take the case of a cylindrical wave and let it be divided into a system of narrow strips by lines drawn on the surface of the wave parallel to the axis of the cylinder as indicated in Art. 148, p. 254. Further, let each of these strips be replaced by its central effective portion, so that the whole wave is reduced to its equatorial band. By this process the calculation of the effect of the wave is reduced to finding that of a circular band, and we know that of the whole band it is only a small portion in the neighbourhood of the pole that is effective. The whole effect is therefore equivalent to that of a small arc of a circle. Let $a$ be the radius of this circle—that is, the radius of the cylindrical wave—and let $b$ be the distance of the point O (Fig. 143) at which the effect is sought, from the pole P of the wave. Then if OM $= r$ and if PM $= s$ we have

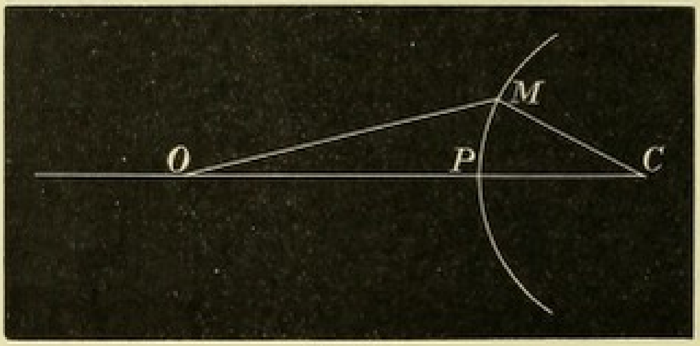


Fig. 143.

$$r^2 = (a+b)^2 + a^2 - 2a(a+b)\cos s/a,$$

consequently when $s$ is small we have approximately

$$r - b = \frac{a+b}{2ab} s^2.$$

That is, the relative path retardation of the vibration from M is proportional to the square of the arc PM. The phase difference is consequently

$$\phi = \frac{2\pi}{\lambda}(r-b) = \frac{\pi(a+b)}{ab\lambda} s^2.$$

Hence if we assume that the amplitude of the vibration produced at O by an element $ds$ of the circle is simply proportional to the length of the element, the quantity $fd\mathrm{S}/r$ in the preceding integrals must be replaced by $ds$, and the expression for the intensity becomes

$$\mathrm{I} = \left[\int \cos\phi ds\right]^2 + \left[\int \sin\phi ds\right]^2,$$

or, writing $\phi$ in the form $\frac{1}{2}\pi v^2$ where $v = s\sqrt{2(a+b)/ab\lambda}$, we have $ds$ proportional to $dv$, and therefore I may be expressed in the form

$$\mathrm{I} = \left[\int \cos \tfrac{1}{2}\pi v^2 dv\right]^2 + \left[\int \sin \tfrac{1}{2}\pi v^2 dv\right]^2.$$

This is the formula deduced by Fresnel, and the two integrals which appear in it are known as Fresnel's integrals.

It should not be lost sight of that in expressing I in this form certain assumptions and approximations have been made, and that it is on account of these assumptions and approximations that integrals of this form appear. Thus it is practically assumed that the effective portion of the equatorial belt is so small that $\phi$ is sensibly proportional to $s^2$, and that $r$ is approximately constant and equal to $b$. Further, the amplitude of the vibration contributed by any element of area is taken proportional to the area, and this amounts to assuming either that $f/r$ is constant, or that the function $f$ is constant as well as $r$.

Now in examining the effect of a plane or a spherical wave at any point we have seen (Arts. 52 and 53) that the only factor left by which the approximate rectilinear propagation may be explained is that the function $f$ is not constant, but is such that it decreases as the obliquity increases—that is, as $r$ increases. Consequently it is not legitimate to assume either that $f$ is constant or that $f/r$ is constant, for it is on the variations of $f$ and $r$ (however small) that the whole outstanding effect depends in the case of a complete wave.

The final result obtained by Fresnel amounts to saying at once that the effective portion of the wave front is confined to a small arc which is so short that throughout it $\phi$ may be taken proportional to $s^2$, and this is what stamps Fresnel's integrals with their peculiar form. This assumption confines the effective portion of the wave to a small portion around the pole, and therefore virtually introduces a law of diminution of effect with obliquity; or, in other words, a form of the function $f$.

**Fortuitous law of obliquity.**

To determine the law which Fresnel introduced inadvertently by these assumptions let us suppose that the wave is divided into ring elements, as in Art. 53, so that $d\mathrm{S}$ is proportional to $rd\phi$. Then $fd\mathrm{S}/r = kfd\phi$ where $k$ is a constant, and $f\cos\phi d\mathrm{S}/r = kf\cos\phi d\phi$. Now in order that this may take the form $\cos z^2 dz$ it is only necessary to suppose that

$$f \propto \frac{1}{\sqrt{\phi}} \propto \frac{1}{\sqrt{r-b}}.$$

Thus the law inadvertently introduced is that $f$ varies inversely as the square root of the phase retardation, and consequently its rate of decrease as we recede from the pole is very rapid, and the effect of a small area surrounding the pole must be approximately the same as that of the whole wave. We should therefore expect that the value of either of Fresnel's integrals when taken between the limits 0 and $v$ should be sensibly the same whether we take $v$ fairly large or infinitely great. In fact, as $v$ increases from zero the value of each integral fluctuates and passes through a series of maxima and minima, but these fluctuations soon become less and less pronounced, the maxima decreasing while the minima increase until they become sensibly equal, and the value of the integral remains sensibly constant as $v$ increases indefinitely.

This may be seen geometrically by plotting the curve $y = \cos x^2$, in which the value of the ordinate varies periodically between the limits $\pm 1$ as $x$ increases (Fig. 144). The area of this curve is $\int y dx = \int \cos x^2 dx$, and may therefore be taken to represent the value of one of Fresnel's integrals. Now while $y$ varies between constant limits $\pm 1$, it is to be remarked that the distance between two consecutive points of intersection of the curve with the axis of $x$ continually diminishes as $x$ increases. In fact, if $x$ and $x + h$ be the abscissæ of two consecutive points of intersection of the curve with the axis of $x$, we have

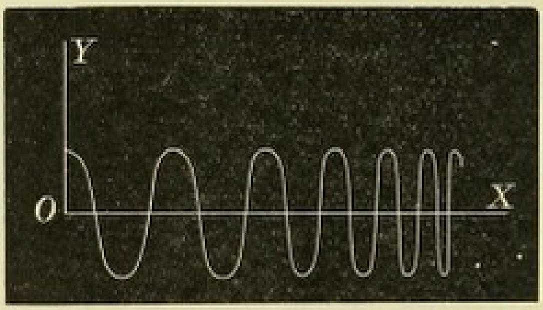


Fig. 144.

$$x^2 = (2n+1)\pi/2, \quad \text{and } (x+h)^2 = (2n+3)\pi/2,$$

so that by subtraction we obtain

$$h^2 + 2xh = \pi,$$

and, therefore, as $x$ increases $h$ must diminish. The total area of the curve is thus the sum of a system of loops which are alternately positive and negative and which decrease indefinitely in absolute value. Hence the integral, after passing through a series of maxima and minima, rapidly attains a stationary value, as shown in the following table after Gilbert :—

TABLE OF FRESNEL'S INTEGRALS (Gilbert)[1]

| $v$ | $\int_0^v \cos \frac{1}{2}\pi v^2 dv$ | $\int_0^v \sin \frac{1}{2}\pi v^2 dv$ | $v$ | $\int_0^v \cos \frac{1}{2}\pi v^2 dv$ | $\int_0^v \sin \frac{1}{2}\pi v^2 dv$ |
|---|---|---|---|---|---|
| 0·0 | 0·0000 | 0·0000 | 2·6 | 0·3389 | 0·5500 |
| 0·1 | 0·0999 | 0·0005 | 2·7 | 0·3926 | 0·4529 |
| 0·2 | 0·1999 | 0·0042 | 2·8 | 0·4675 | 0·3915 |
| 0·3 | 0·2994 | 0·0141 | 2·9 | 0·5624 | 0·4102 |
| 0·4 | 0·3975 | 0·0334 | 3·0 | 0·6057 | 0·4963 |
| 0·5 | 0·4923 | 0·0647 | 3·1 | 0·5616 | 0·5818 |
| 0·6 | 0·5811 | 0·1105 | 3·2 | 0·4663 | 0·5933 |
| 0·7 | 0·6597 | 0·1721 | 3·3 | 0·4057 | 0·5193 |
| 0·8 | 0·7230 | 0·2493 | 3·4 | 0·4385 | 0·4297 |
| 0·9 | 0·7648 | 0·3398 | 3·5 | 0·5326 | 0·4153 |
| 1·0 | 0·7799 | 0·4383 | 3·6 | 0·5880 | 0·4923 |
| 1·1 | 0·7638 | 0·5365 | 3·7 | 0·5419 | 0·5750 |
| 1·2 | 0·7154 | 0·6234 | 3·8 | 0·4481 | 0·5656 |
| 1·3 | 0·6386 | 0·6863 | 3·9 | 0·4223 | 0·4752 |
| 1·4 | 0·5431 | 0·7135 | 4·0 | 0·4984 | 0·4205 |
| 1·5 | 0·4453 | 0·6975 | 4·1 | 0·5737 | 0·4758 |
| 1·6 | 0·3655 | 0·6383 | 4·2 | 0·5417 | 0·5632 |
| 1·7 | 0·3238 | 0·5492 | 4·3 | 0·4494 | 0·5540 |
| 1·8 | 0·3363 | 0·4509 | 4·4 | 0·4383 | 0·4623 |
| 1·9 | 0·3945 | 0·3734 | 4·5 | 0·5258 | 0·4342 |
| 2·0 | 0·4883 | 0·3434 | 4·6 | 0·5672 | 0·5162 |
| 2·1 | 0·5814 | 0·3743 | 4·7 | 0·4914 | 0·5669 |
| 2·2 | 0·6362 | 0·4556 | 4·8 | 0·4338 | 0·4968 |
| 2·3 | 0·6268 | 0·5525 | 4·9 | 0·5002 | 0·4351 |
| 2·4 | 0·5550 | 0·6197 | 5·0 | 0·5636 | 0·4992 |
| 2·5 | 0·4574 | 0·6192 | $\infty$ | 0·5000 | 0·5000 |

When the values of the integrals expressing the intensity are known, then if one of them be denoted by $x$, and the corresponding value of the other by $y$, it is clear that when a curve is constructed with $x$ and $y$ as co-ordinates, the radius vector $r$ drawn from the origin to any point on the curve will represent the resultant vibration arising from the corresponding portion of the wave. For we have $r^2 = x^2 + y^2$, and consequently $r^2$ represents the intensity of the illumination, or $r$ represents the amplitude of the resultant vibration and the angle it makes with the axis of $x$ represents its phase. The curve constructed in this manner is, in fact, the vibration spiral. Using the values of Fresnel's integrals given in the foregoing table, M. Cornu constructed the spiral shown in Fig. 145. The branch $OM_1M_2J$ refers to one half of the wave, and the branch OJ′ to the other. The part $OM_1$ represents the first half-period element, $M_1M_2$ the second, $M_2M_3$ the third, and so on, the successive convolutions winding with ever increasing curvature round a central point J. This point is situated

[1] Gilbert, *Mém. couronnés de l'Acad. de Bruxelles*, tom. xxxi. p. 1, 1863.

on the line bisecting the angle between the axis, for its co-ordinates are

$$x=\int_0^\infty \cos \tfrac{1}{2}\pi v^2 dv = \tfrac{1}{2}, \quad \text{and } y=\int_0^\infty \sin \tfrac{1}{2}\pi v^2 dv = \tfrac{1}{2},$$

consequently at J and J′ we have $x = y$—that is, these points lie on the line bisecting the angle XOY. The phase of the resultant would thus appear to be only 45° in advance of that which arrives from the central element. It must be remembered, however, that in evaluating I by this method the wave surface has been reduced by the strip

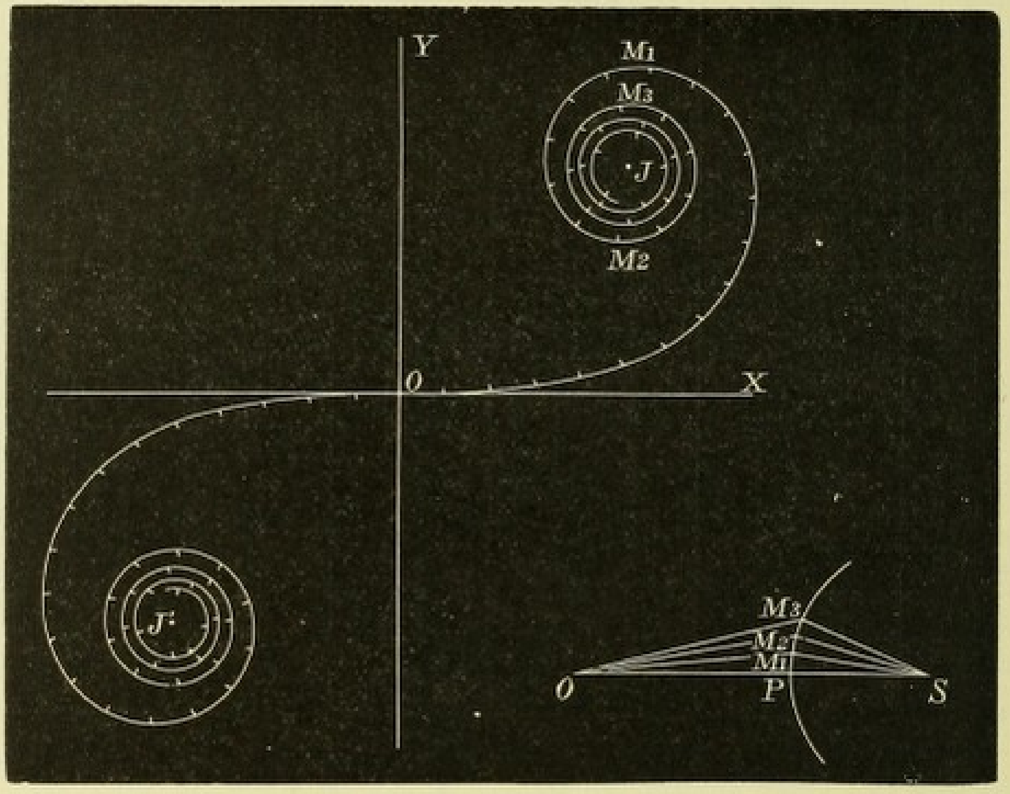



Fig. 145.—Cornu's Spiral.

method to an equatorial band, so that each element of this band is the resultant of a strip. The phase of each element of the band is thereby advanced by 45° relatively to that of the vibration from the pole of the strip. The whole phase of the resultant is consequently 90° in advance of that arriving from the pole of the wave surface. This rotates the spiral through 45° and throws the points J and J′ on the axis OY, as in Fig. 127.

Methods of evaluating Fresnel's integrals have been given by Fresnel, Gilbert, Cauchy, and Knochenhauer. These methods are noticed in the following examples and will be found at length in Verdet's *Œuvres*, tom. v. p. 328, etc.; *Optique Physique*, tom. i.

### *Examples*

1. At what distance from a slit of width $2c$ is it necessary to place a screen so that the central fringe may be of minimum intensity? and determine this intensity (Cornu).

Here we must determine the points of the two branches of the spiral which are nearest. Since in this case they are symmetrically situated, the line joining them passes through $O$, and with a compass the minimum distance is found to correspond to $v = \pm 1{\cdot}875$. Since $s = c$ we have

$$\frac{2(a+b)}{ab\lambda}c^2 = (1{\cdot}875)^2.$$

Again the intensity is measured by the square of the chord of the curve, and this is $(1{\cdot}08)^2 = 1{\cdot}17$ where the intensity of the light due to the whole wave is $\mathrm{JJ}'^2 = 2$. The relative intensity of the central band is therefore $0{\cdot}585$.

2. Prove that

$$\int_i^{i+u} \cos \tfrac{1}{2}\pi v^2 dv = \frac{1}{\pi i}\left[\sin \tfrac{1}{2}\pi(i^2 + 2iu) - \sin \tfrac{1}{2}\pi i^2\right],$$

$$\int_i^{i+u} \sin \tfrac{1}{2}\pi v^2 dv = \frac{1}{\pi i}\left[-\cos \tfrac{1}{2}\pi(i^2 + 2iu) + \cos \tfrac{1}{2}\pi i^2\right]$$

(Fresnel, *Œuvres*, tom. i. p. 319),

where $u$ is a small quantity and $i$ given.

[Replacing $v$ by $i + u$ we have

$$\cos \tfrac{1}{2}\pi v^2 = \cos \tfrac{1}{2}\pi(i+u)^2 = \cos \tfrac{1}{2}\pi(i^2 + 2iu)$$

neglecting $u^2$. Hence

$$\int_i^{i+u} \cos \tfrac{1}{2}\pi v^2 dv = \int_o^u \cos \tfrac{1}{2}\pi(i^2 + 2iu)du$$

$$= \cos \tfrac{1}{2}\pi i^2 \int_o^u \cos \pi iu du - \sin \tfrac{1}{2}\pi i^2 \int_o^u \sin \pi iu du,$$

which is integrable at once.

By this method Fresnel calculated the values of the integrals, taking $u = 0 . 1$ and $i$ successively equal to 0, 0 . 1, 0 . 2, 0 . 3, etc.]

3. Using the equations of Ex. 2, show that

$$\left[\int_i^{i+u} \cos \tfrac{1}{2}\pi v^2 dv\right]^2 + \left[\int_i^{i+u} \sin \tfrac{1}{2}\pi v^2 dv\right]^2 = \frac{4}{\pi^2 i^2}\sin^2 \tfrac{1}{2}\pi iu.$$

4. Prove that

$$\int_o^v \cos \tfrac{1}{2}\pi v^2 dv = \cos \tfrac{1}{2}\pi v^2\left(\frac{v}{1} - \frac{\pi^2 v^5}{1\,.\,3\,.\,5} + \frac{\pi^4 v^9}{1\,.\,3\,.\,5\,.\,7\,.\,9} - \dots\right)$$

$$+ \sin \tfrac{1}{2}\pi v^2\left(\frac{\pi v^3}{1\,.\,3} - \frac{\pi^3 v^7}{1\,.\,3\,.\,5\,.\,7} + \frac{\pi^5 v^{11}}{1\,.\,3\,.\,5\,.\,7\,.\,9\,.\,11} - \dots\right)$$

(Knochenhauer, *Die Undulationstheorie des Lichtes*, p. 36).

[Integrating by parts, we have

$$\int_o^v \cos \tfrac{1}{2}\pi v^2 dv = v \cos \tfrac{1}{2}\pi v^2 + \pi \int_o^v v^2 \sin \tfrac{1}{2}\pi v^2 dv,$$

$$\int_o^v v^2 \sin \tfrac{1}{2}\pi v^2 dv = \frac{v^3}{3}\sin \tfrac{1}{2}\pi v^2 - \frac{\pi}{3}\int_o^v v^4 \cos \tfrac{1}{2}\pi v^2 dv,$$

$$\int_o^v v^4 \cos \tfrac{1}{2}\pi v^2 dv = \frac{v^5}{5}\cos \tfrac{1}{2}\pi v^2 + \frac{\pi}{5}\int_o^v v^6 \sin \tfrac{1}{2}\pi v^2 dv, \text{ etc.]}$$

5. Writing the result of Ex. 4 in the form

$$\int_0^v \cos \tfrac{1}{2}\pi v^2 dv = \mathrm{M} \cos \tfrac{1}{2}\pi v^2 + \mathrm{N} \sin \tfrac{1}{2}\pi v^2,$$

show that

$$\int_0^v \sin \tfrac{1}{2}\pi v^2 dv = \mathrm{M} \sin \tfrac{1}{2}\pi v^2 - \mathrm{N} \cos \tfrac{1}{2}\pi v^2.$$

Hence

$$\left[\int_0^v \cos \tfrac{1}{2}\pi v^2 dv\right]^2 + \left[\int_0^v \sin \tfrac{1}{2}\pi v^2 dv\right]^2 = \mathrm{M}^2 + \mathrm{N}^2,$$

and

$$\frac{d\mathrm{M}}{dv} = 1 - \pi v \mathrm{N}, \quad \frac{d\mathrm{N}}{dv} = \pi v \mathrm{M}.$$

6. Prove that

$$\int_v^\infty \cos \tfrac{1}{2}\pi v^2 dv = \cos \tfrac{1}{2}\pi v^2 \left( \frac{1}{\pi^2 v^3} - \frac{1\,.\,3\,.\,5}{\pi^4 v^7} + \frac{1\,.\,3\,.\,5\,.\,7\,.\,9}{\pi^6 v^{11}} - \ldots \right)$$

$$- \sin \tfrac{1}{2}\pi v^2 \left( \frac{1}{\pi v} - \frac{1\,.\,3}{\pi^3 v^5} + \frac{1\,.\,3\,.\,5\,.\,7}{\pi^5 v^9} + \ldots \right)$$

(Cauchy, *Comptes Rendus*, tom. xv. pp. 534, 573).

[Integrating by parts, we have

$$\int_v^\infty \cos \tfrac{1}{2}\pi v^2 dv = \Big|_v^\infty \frac{1}{\pi v} \sin \tfrac{1}{2}\pi v^2 + \int_v^\infty \frac{1}{\pi v^2} \sin \tfrac{1}{2}\pi v^2 dv,$$

$$\int_v^\infty \frac{1}{\pi v^2} \sin \tfrac{1}{2}\pi v^2 dv = -\Big|_v^\infty \frac{1}{\pi^2 v^3} \cos \tfrac{1}{2}\pi v^2 - 3\int_v^\infty \frac{1}{\pi^2 v^4} \cos \tfrac{1}{2}\pi v^2 dv, \text{ etc.}]$$

7. Writing the equation of Ex. 6 in the form

$$\int_v^\infty \cos \tfrac{1}{2}\pi v^2 dv = \mathrm{P} \cos \tfrac{1}{2}\pi v^2 - \mathrm{Q} \sin \tfrac{1}{2}\pi v^2,$$

show that

$$\int_v^\infty \sin \tfrac{1}{2}\pi v^2 dv = \mathrm{P} \sin \tfrac{1}{2}\pi v^2 + \mathrm{Q} \cos \tfrac{1}{2}\pi v^2,$$

and also

$$\frac{d\mathrm{P}}{dv} = \pi v \mathrm{Q} - 1, \quad \frac{d\mathrm{Q}}{dv} = -\pi v \mathrm{P}.$$

8. Using the same notation, show that

$$\int_0^v \cos \tfrac{1}{2}\pi v^2 dv = \tfrac{1}{2} - \mathrm{P} \cos \tfrac{1}{2}\pi v^2 + \mathrm{Q} \sin \tfrac{1}{2}\pi v^2,$$

$$\int_0^v \sin \tfrac{1}{2}\pi v^2 dv = \tfrac{1}{2} - \mathrm{P} \sin \tfrac{1}{2}\pi v^2 - \mathrm{Q} \cos \tfrac{1}{2}\pi v^2.$$

[We have

$$\int_0^v \cos \tfrac{1}{2}\pi v^2 dv = \int_0^\infty \cos \tfrac{1}{2}\pi v^2 dv - \int_v^\infty \cos \tfrac{1}{2}\pi v^2 dv = \tfrac{1}{2} - \mathrm{P} \cos \tfrac{1}{2}\pi v^2 + \mathrm{Q} \sin \tfrac{1}{2}\pi v^2$$

by Ex. 7.]

9. Prove that

$$\int_0^v \cos \tfrac{1}{2}\pi v^2 dv = \tfrac{1}{2} - \mathrm{G} \cos \tfrac{1}{2}\pi v^2 + \mathrm{H} \sin \tfrac{1}{2}\pi v^2,$$

$$\int_0^v \sin \tfrac{1}{2}\pi v^2 dv = \tfrac{1}{2} - \mathrm{G} \sin \tfrac{1}{2}\pi v^2 - \mathrm{H} \cos \tfrac{1}{2}\pi v^2,$$

where

$$H=\frac{1}{\pi\sqrt{2}}\int_0^\infty \frac{e^{-ux}dx}{\sqrt{x}(1+x^2)},\quad G=\frac{1}{\pi\sqrt{2}}\int_0^\infty \frac{e^{-ux}\sqrt{x}dx}{1+x^2}$$

(Gilbert, *Mém. couronnés de l'Acad. de Bruxelles*, tom. xxxi. p. 1).

[If $u=\frac{1}{2}\pi v^2$, we have

$$v=\sqrt{\frac{2u}{\pi}},\quad \text{and } dv=\frac{du}{\sqrt{2\pi u}},$$

therefore

$$\int_0^v \cos\tfrac{1}{2}\pi v^2 dv=\frac{1}{\sqrt{2\pi}}\int_0^u \cos u\frac{du}{\sqrt{u}}.$$

But

$$\int_0^\infty e^{-z^2}dz=\frac{\sqrt{\pi}}{2}.$$

Hence, by writing $z^2=ux$, and regarding $u$ as a constant, we find

$$\int_0^\infty \frac{e^{-ux}dx}{\sqrt{x}}=\sqrt{\frac{\pi}{u}},\quad \therefore \frac{1}{\sqrt{u}}=\frac{1}{\sqrt{\pi}}\int_0^\infty \frac{e^{-ux}dx}{\sqrt{x}},$$

and we obtain

$$\int_0^v \cos\tfrac{1}{2}\pi v^2 dv=\frac{1}{\sqrt{2\pi}}\int_0^u \frac{\cos u du}{\sqrt{\pi}}\int_0^\infty \frac{e^{-ux}dx}{\sqrt{x}},$$

$$=\frac{1}{\pi\sqrt{2}}\int_0^\infty \frac{dx}{\sqrt{x}}\int_0^u e^{-ux}\cos u du.$$

The integration with respect to $u$ is easily effected by parts, thus

$$\int_0^u e^{-ux}\cos u du=\frac{x}{1+x^2}-\frac{e^{-ux}(x\cos u-\sin u)}{1+x^2},$$

and the required integral takes the form

$$\int_0^v \cos\tfrac{1}{2}\pi v^2 dv=\frac{1}{\pi\sqrt{2}}\left[\int_0^\infty \frac{\sqrt{x}dx}{1+x^2}-\cos u\int_0^\infty \frac{e^{-ux}\sqrt{x}dx}{1+x^2}+\sin u\int_0^\infty \frac{e^{-ux}dx}{\sqrt{x}(1+x^2)}\right].$$

The value of the first integral in the bracket is easily found to be $\pi/\sqrt{2}$, by substituting $z^2$ for $x$, and we have finally

$$\int_0^v \cos\tfrac{1}{2}\pi v^2 dv=\tfrac{1}{2}-G\cos u+H\sin u$$

as required.]

Comparing these expressions with those of Cauchy (Ex. 8) we have

$$P=G,\quad Q=H.$$

10. If

$$C=\int_0^v \cos\tfrac{1}{2}\pi v^2 dv,$$

and

$$S=\int_0^v \sin\tfrac{1}{2}\pi v^2 dv,$$

prove that

$$G = \tfrac{1}{2}(\cos u + \sin u) - C\cos u - S\sin u,$$
$$H = \tfrac{1}{2}(\cos u - \sin u) + C\sin u - S\cos u$$

where, as before, $u = \frac{1}{2}\pi v^2$.

[Solve from Ex. 9.]

11. Apply Ex. 5 to prove that

$$G = \tfrac{1}{2}(\cos u + \sin u) - M,$$
$$H = \tfrac{1}{2}(\cos u - \sin u) + N.$$

12. If $u = 0$, prove that

$$G = H = \tfrac{1}{2}.$$

13. If axes of reference be taken in the wave front at P, or parallel to it through O, we have

$$\rho^2 = x^2 + y^2, \quad \text{and } dS = dxdy.$$

Hence if $f$ = constant and $r = b$, that is, if the effective portion is limited to a small area around the pole P, we have $\delta = \pi\rho^2/b\lambda$ (Ex. 4, p. 274), and the intensity of the illumination is proportional to

$$[\int\cos\tfrac{1}{2}\pi c(x^2 + y^2)dxdy]^2 + [\int\sin\tfrac{1}{2}\pi c(x^2 + y^2)dxdy]^2,$$

where $c$ is a constant $(2/b\lambda)$ in the case of a plane wave, and it represents $2(a+b)/ab\lambda$ for a spherical wave of radius $a$.

Denoting these integrals by M and N respectively, we have by expansion

$$M = \int\cos\tfrac{1}{2}\pi cx^2dx\int\cos\tfrac{1}{2}\pi cy^2dy - \int\sin\tfrac{1}{2}\pi cx^2dx\int\sin\tfrac{1}{2}\pi cy^2dy,$$
$$N = \int\sin\tfrac{1}{2}\pi cx^2dx\int\cos\tfrac{1}{2}\pi cy^2dy + \int\cos\tfrac{1}{2}\pi cx^2dx\int\sin\tfrac{1}{2}\pi cy^2dy.$$

Replacing $cx^2$ or $cy^2$ by $v^2$, the calculation of these integrals is reduced to the calculation of the integrals (Fresnel's)

$$C = \int\cos\tfrac{1}{2}\pi v^2dv, \quad S = \int\sin\tfrac{1}{2}\pi v^2dv.$$

## *Particular Cases*

(*a*) *Complete Wave.*—If the wave is unobstructed, we have

$$C = \int_{-\infty}^{+\infty}\cos\tfrac{1}{2}\pi v^2dv = 1, \quad S = \int_{-\infty}^{+\infty}\sin\tfrac{1}{2}\pi v^2dv = 1.$$

Hence

$$M = 0, \quad \text{and } N = \frac{2}{c}.$$

Therefore

$$M^2 + N^2 = \frac{4}{c^2},$$

and

$$\tan\phi = N/M = \infty, \quad \therefore \phi = \tfrac{1}{2}\pi.$$

(*b*) *Straight Edge.*—If the wave passes over the edge of an opaque obstacle, we integrate with respect to $y$ between the limits $+\infty$ and $-\infty$, so that

$$M=\int\cos\tfrac{1}{2}\pi v^2dv-\int\sin\tfrac{1}{2}\pi v^2dv,$$

$$N=\int\cos\tfrac{1}{2}\pi v^2dv+\int\sin\tfrac{1}{2}\pi v^2dv,$$

$$M^2+N^2=2[\int\cos\tfrac{1}{2}\pi v^2dv]^2+2[\int\sin\tfrac{1}{2}\pi v^2dv]^2.$$

Outside the shadow the integration extends over one half of the wave and part of the other half, thus

$$\int_o^\infty\cos\tfrac{1}{2}\pi v^2dv+\int_o^v\cos\tfrac{1}{2}\pi v^2dv=\tfrac{1}{2}+\int_o^v\cos\tfrac{1}{2}\pi v^2dv.$$

Hence outside

$$I=\left[\tfrac{1}{2}+\int_o^v\cos\tfrac{1}{2}\pi v^2dv\right]^2+\left[\tfrac{1}{2}+\int_o^v\sin\tfrac{1}{2}\pi v^2dv\right]^2.$$

But inside the shadow the integration extends over part of half a wave, from $v$ to $\infty$, and

$$\int_v^\infty\cos\tfrac{1}{2}\pi v^2dv=\int_o^\infty\cos\tfrac{1}{2}\pi v^2dv-\int_o^v\cos\tfrac{1}{2}\pi v^2dv=\tfrac{1}{2}-\int_o^v\cos\tfrac{1}{2}\pi v^2dv.$$

Hence inside the shadow

$$I=\left[\tfrac{1}{2}-\int_o^v\cos\tfrac{1}{2}\pi v^2dv\right]^2+\left[\tfrac{1}{2}-\int_o^v\sin\tfrac{1}{2}\pi v^2dv\right]^2.$$

The cases of a narrow wire and narrow slit are treated in the same manner.

**162. On the Scattering Action of very Small Particles, and the Colour of Skylight.**—In applying the wave theory to deduce the ordinary laws of reflection of light (chap. iv.) the linear dimensions of the obstacle at which reflection occurs have been supposed enormously great compared with the length of the reflected waves. For this reason the ordinary laws of reflection involve no reference to the wave length, and the screening action of the obstacle is perfect except for a narrow band around the edge of the geometrical shadow. The width of this band, however, and the intensity at any point of it, depend on the wave length (being greater for the longer waves than for the shorter), and consequently a small obstacle will not be so effective as a screen to the longer waves as to the shorter.

Thus we have seen (chap. iii.) that an obstacle will screen a point from the influence of a wave when it is wide enough to cover a large number of half-period elements of the wave front, but the width of any half-period element increases with the wave length, and consequently a given obstacle will act most effectively as a barrier to the waves of shortest length. We are consequently prepared to admit that when ordinary white light passes through a medium (such as the atmosphere) in which very small particles are suspended, the waves of greater length will be more freely transmitted than those of higher refrangibility. After passing through a certain thickness of such a medium the light from a white source should consequently appear yellowish (like the snow on a distant mountain), and as the thickness increases the tint should become reddish.

To say that the longer waves are most freely transmitted is equivalent to saying that the shorter are most copiously reflected, and we should therefore expect the light reflected (or scattered) in any direction by such particles of matter to be rich in the rays of higher refrangibility. Now the light which reaches us from the open sky is that part of the light of the sun that has been reflected, or scattered, by the fine particles of matter suspended in the atmosphere, and if these particles are small compared with the wave length of light, the tint of sky should belong to the blue end of the spectrum rather than to the red.

In estimating the quality of this light, however, it must be remembered that it has suffered from the modifying action of transmission as well as scattering. For it is clear that the light which reaches the eye after scattering in a certain locality is merely the residuum of the solar light which has survived after transmission through a certain thickness of the air, as well as the scattering action of the particles. Thus the light first passes through a certain thickness of air and is modified by transmission. This modified light is then reflected or scattered by certain particles, and the scattered light is subsequently transmitted through some thickness of air to the eye of the observer. Now we have seen that transmission is detrimental to the rays of higher refrangibility, while the scattering cuts off those of lower refrangibility, and the light which survives and reaches the eye will therefore be weak in both ends of the spectrum. In other words, it will be composed chiefly of the waves of intermediate length from the region of the blue or green. The exact colour of course will depend on the size and plenitude of the particles—for example, when the particles are relatively large they will affect all wave lengths of light in practically the same degree, and the scattered light will be white.

This explanation of the blue colour of the sky was proposed by Lord Rayleigh[1] in 1871, and the law according to which the scattering takes place for waves of different lengths may be easily deduced from elementary considerations. Thus we have seen (Art. 159) that if the amplitude of the vibration contributed by an element of a wave

[1] The Hon. J. W. Strutt, *Phil. Mag.* vol. xli. pp. 107, 275.

surface to any point be taken to vary directly as the element of area, and inversely as the distance, then the complete expression for the amplitude must contain $\lambda^{-1}$, so that the intensity must vary inversely as the square of the wave length. In the same way if we consider the light scattered in any portion of space laden with very small particles, and if we assume that the amplitude of the vibration contributed by any element of volume $dV$ of this space is proportional to that element of volume, and inversely as the distance, then in the expression of Art. 159, $dS$ must be replaced by $dV$, and the complete expression for the amplitude must contain $\lambda^{-2}$, so that the intensity of the scattered light will vary inversely as the fourth power of $\lambda$—that is, when scattering alone is considered,

$$I_s \propto \frac{1}{\lambda^4}.$$

Now the light which reaches the eye is scattered, and also transmitted (both before and after scattering) through some thickness $x$ of the medium. Hence its composition will be determined by finding the effect of transmission on the quantity $I_s$. This problem is equivalent to that of finding the intensity of a pencil of light after transmission through an absorbing medium when the absorption varies inversely as the fourth power of the wave length. Hence, as in Art. 282, we have

$$\frac{dI}{I} = -\frac{kdx}{\lambda^4},$$

where $k$ is a constant, and $dI$ the change of intensity in passing through a layer of thickness $dx$. The intensity of the light reaching the eye after suffering scattering as well as transmission through a total thickness $x$ of the medium is therefore

$$I = I_s e^{-kx/\lambda^4},$$

and substituting for $I_s$ we have finally

$$I = \frac{A}{\lambda^4} e^{-kx/\lambda^4}.$$

This expression exhibits the joint effects of scattering and transmission, and shows how I diminishes for the large values of $\lambda$ as well as for the small. The maximum value of I corresponds to some intermediate wave length $\lambda_m$ given by the equation

$$\lambda_m^{\ 4} = kx,$$

and the corresponding maximum value of I is

$$I_m = A\lambda_m^{-4}e^{-1} = A/ekx,$$

while the intensity I, corresponding to any wave length $\lambda$, is related to $I_m$ by the equation

$$\frac{I}{I_m} = \frac{\lambda_m^4}{\lambda^4} e^{1-\frac{\lambda_m^4}{\lambda^4}}.$$

In order to test this theory Lord Rayleigh compared the blue light of the sky taken from the neighbourhood of the zenith with sunlight diffused through white paper. The results of this comparison are contained in the following table for the fixed lines C, D, $b_3$, F of the spectrum, and, considering the difficulties and uncertainties of the comparison, the observed and calculated values appear to agree tolerably well. It must of course be observed that the comparison is based on the assumption that the light diffused through white paper is similar to the scattered light which illuminates the sky.

| C. | D. | $b_3$. | F. | |
|---|---|---|---|---|
| 25 | 40 | 63 | 80 | (calculated) |
| 25 | 41 | 70 | 90 | (observed) |

It appears therefore that the skylight, when compared with that diffused through white paper, was bluer than that required by the theory; and this, Lord Rayleigh suggests, may arise from the possible yellowness of the paper, or from the yellowness of the sunlight when it reaches us compared with that in the upper regions of the atmosphere where it is diffused.

**163. Diffraction in Optical Instruments—Circular Aperture.**— In many optical instruments the aperture through which the light enters is limited by a circular stop, and for this reason the study of diffraction in the case of a circular aperture has attracted special attention.

For example—when a point source of light, such as a distant star, is viewed through a telescope, the wave falling upon the object glass is limited by a circular aperture, and instead of having a point image of the source (as the geometrical theory would lead us to expect with an aplanatic lens) in the focal plane of the telescope, there is a diffraction pattern [1] which consists of a bright central spot surrounded by a series of rings alternately bright and dark. With white light the central spot is approximately white and the rings are coloured. The

[1] Noticed by W. Herschel in 1782, *Phil. Trans. Roy. Soc.* p. 52, 1805.

brilliancy of the rings diminishes rapidly from the centre outwards. For the exhibition of several rings a strong source of light is required, and with a faint source only the central spot can be observed. The diameter of the central spot in all cases varies inversely as the diameter of the aperture.

Let the incident light be parallel and fall normally on the aperture, and let the diffracted pencil make an angle $\theta$ with the normal to its plane. Let C be the central point of the aperture, CN a normal to its plane, and CO the direction of the diffracted beam, then $\mathrm{OCN}=\theta$, and the plane of diffraction at C—that is, the plane of CO and CN—meets the plane of the aperture in a line AB. Let P be any point of the aperture, and let the radius vector $\mathrm{CP}(=\rho)$ make an angle $\phi$ with AC. Then the element of area at P is

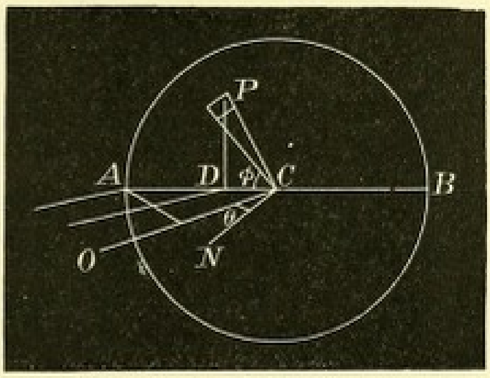


Fig. 146.

$$dS=\rho d\phi d\rho,$$

and the phase retardation $\delta$ of the vibration from this element, relative to that from A, is the same as for that from D, where PD is perpendicular to AB, and therefore

$$\delta=\frac{2\pi}{\lambda}\mathrm{AD}\sin\theta=\frac{2\pi}{\lambda}\sin\theta\,.\,(r-\rho\cos\phi)=h(r-\rho\cos\phi),$$

where $h=\frac{2\pi}{\lambda}\sin\theta$, and $r$ is the radius of the aperture. Hence the vibration excited by the element $dS$ at P is proportional to

$$\sin\{\omega t+h(r-\rho\cos\phi)\}\rho d\phi d\rho,$$

and the resultant vibration for the complete aperture will be

$$\int_0^{2\pi}\int_0^r\rho\sin(\omega t+hr-h\rho\cos\phi)d\phi d\rho$$

$$=\sin(\omega t+hr)\int_0^{2\pi}\int_0^r\rho\cos(h\rho\cos\phi)d\phi d\rho-\cos(\omega t+hr)\int_0^{2\pi}\int_0^r\rho\sin(h\rho\cos\phi)d\phi d\rho.$$

The final double integral in this expression is zero, for the elements of it which arise from any two points situated at equal distances on opposite sides of C are of opposite signs and destroy each other. Hence the resultant vibration is

$$\sin(\omega t+hr)\int_0^{2\pi}\int_0^r\rho\cos(h\rho\cos\phi)d\phi d\rho,$$

and the intensity in the direction CO is

$$I=\left[\int_0^{2\pi}\int_0^r \rho\cos(h\rho\cos\phi)d\phi d\rho\right]^2.$$

The integration with respect to $\rho$ is obtained at once by parts, thus—

$$\int_0^r \rho\cos(h\rho\cos\phi)d\rho=\left\{{}_0^r\frac{\rho}{h\cos\phi}\sin(h\rho\cos\phi)\right\}-\frac{1}{h\cos\phi}\int_0^r\sin(h\rho\cos\phi)d\rho$$

$$=\frac{r}{h\cos\phi}\sin(hr\cos\phi)+\frac{1}{h^2\cos^2\phi}\left\{\cos(hr\cos\phi)-1\right\}$$

$$=r^2\frac{\sin(hr\cos\phi)}{(hr\cos\phi)}-\tfrac{1}{2}r^2\frac{\sin^2(\frac{1}{2}hr\cos\phi)}{(\frac{1}{2}hr\cos\phi)^2}.$$

Hence, writing $2m=hr=\frac{2\pi r}{\lambda}\sin\theta$, we have

$$I=\left[r^2\int_0^{2\pi}\frac{\sin(2m\cos\phi)}{(2m\cos\phi)}d\phi-\tfrac{1}{2}r^2\int_0^{2\pi}\frac{\sin^2(m\cos\phi)}{(m\cos\phi)^2}d\phi\right]^2.$$

The further integration, with respect to $\phi$, may be obtained in series, for we have

$$\frac{\sin x}{x}=1-\frac{x^2}{\lfloor 3}+\frac{x^4}{\lfloor 5}-\frac{x^6}{\lfloor 7}+\dots$$

$$\frac{\sin^2 x}{x^2}=\frac{1-\cos 2x}{2x^2}=1-\frac{2^3x^2}{\lfloor 4}+\frac{2^5x^4}{\lfloor 6}-\frac{2^7x^6}{\lfloor 8}+\dots$$

Therefore

$$\sqrt{I}=r^2\int_0^{2\pi}\left\{(1-\frac{(2m\cos\phi)^2}{\lfloor 3}+\frac{(2m\cos\phi)^4}{\lfloor 5}\dots\right\}d\phi$$

$$-\tfrac{1}{2}r^2\int_0^{2\pi}\left\{1-\frac{2^3(m\cos\phi)^2}{\lfloor 4}+\frac{2^5(m\cos\phi)^4}{\lfloor 6}\dots\right\}d\phi.$$

But we have

$$\int_0^{2\pi}\cos^{2n}x\,dx=\frac{1.3.5\dots(2n-1)}{2.4.6\dots 2n}2\pi,$$

therefore finally

$$\sqrt{I}=\pi r^2\left\{1-\frac{1}{2}\left(\frac{m}{1}\right)^2+\frac{1}{3}\left(\frac{m^2}{\lfloor 2}\right)^2-\frac{1}{4}\left(\frac{m^3}{\lfloor 3}\right)^2+\frac{1}{5}\left(\frac{m^4}{\lfloor 4}\right)^2\dots\right\}.$$

Or denoting the series within the bracket by S, we have

$$I=(\pi r^2)^2S^2,$$

which is the result obtained by Airy.[1]

The series S is convergent for all values of $m$ and passes alternately

[1] G. B. Airy, *Camb. Phil. Trans.* p. 283, 1834.

through positive and negative values as $m$ increases from zero. The intensity consequently presents maximum and zero values, corresponding to $\frac{dS}{dm}=0$, and $S=0$, respectively; and when the corresponding values of $m$ have been found, the deviation $\theta$ is given by the equation

$$m=\frac{\pi r \sin\theta}{\lambda}, \quad \text{or } \sin\theta=\frac{m\lambda}{\pi r}.$$

Hence the deviation corresponding to any bright or dark ring, being small, is proportional to $\lambda$ and inversely as the radius of the aperture.

The subjoined table taken from Verdet's *Optique Physique* contains the values of $m/\pi$ corresponding to the first few maxima and minima. The table shows the rapidity with which the maxima decrease, and also that the difference between two consecutive values of $m$ tends to become constant and equal to $\frac{1}{2}\pi$.

| | $m/\pi$ | Intensity $S^2$. | | $m/\pi$ | Intensity $S^2$. |
|---|---|---|---|---|---|
| 1st max. | 0 | 1 | 1st min. | 0·610 | 0 |
| 2nd max. | 0·819 | 0·01745 | 2nd min. | 1·116 | 0 |
| 3rd max. | 1·333 | 0·00415 | 3rd min. | 1·619 | 0 |
| 4th max. | 1·847 | 0·00165 | 4th min. | 2·120 | 0 |
| 5th max. | 2·361 | 0·00078 | 5th min. | 2·621 | 0 |

In the foregoing equation $\theta$ is obviously the angular width of a ring as seen from the optical centre, and as this is very small we may write $\theta$ for $\sin\theta$. Hence the angular width of the first dark ring is given by the equation

$$\theta=0{\cdot}61\frac{\lambda}{r}.$$

When two very close point-sources of light are viewed through a telescope their diffraction patterns will overlap, and they cannot be distinguished as distinct sources if the overlapping of the central spots of their images exceeds a certain limit. Now the intensity falls from unity at the centre of the spot to 0·37 at a distance from the centre equal to half the radius of the first dark ring. Consequently if the distance between the centres of the two central spots is equal to the radius of the first dark ring the intensity should be 0·74 half-way between the two centres and unity at each centre. This variation of intensity should be observable, and a double star should therefore be resolved by a telescope when the angular separation of the components is 0·61 $\lambda/r$.

## *Examples*

1. Taking AB (Fig. 146) and a perpendicular to it as axes of $x$ and $y$ respectively, show that with an aperture of any form, the intensity at O in the direction CO making an angle $\theta$ with the wave normal is

$$I=\left[\int y\sin\left(\frac{2\pi}{\lambda}x\sin\theta\right)dx\right]^2+\left[\int y\cos\left(\frac{2\pi}{\lambda}x\sin\theta\right)dx\right]^2.$$

[Divide the aperture into narrow strips perpendicular to the axis of $x$. The area of one of these strips is $ydx$, the length of the strip being $y$ and its width $dx$. Also if the phase of the vibration from the origin of co-ordinates be $\omega t$, the phase retardation of the vibration from the element $ydx$ will be $\frac{2\pi}{\lambda}x\sin\theta$, and hence it will be represented by

$$y\sin\left(\omega t+\frac{2\pi}{\lambda}x\sin\theta\right)dx.$$

The resultant of the whole aperture will be

$$\int y\sin\left(\omega t+\frac{2\pi}{\lambda}x\sin\theta\right)dx$$
$$=\sin\omega t\int y\cos\left(\frac{2\pi}{\lambda}x\sin\theta\right)dx+\cos\omega t\int y\sin\left(\frac{2\pi}{\lambda}x\sin\theta\right)dx.$$

Therefore, etc.]

2. The sine of the deviation ($\sin\theta$) corresponding to a maximum or minimum intensity (Ex. 1) is proportional to the wave length $\lambda$. Hence the linear dimensions of the patterns produced by differently coloured homogeneous lights are proportional to the wave lengths.

3. With similar apertures the sines of the deviations ($\sin\theta$) corresponding to a maximum or minimum intensity of a given order are inversely as the linear dimensions of the apertures.

[For let $m$ be the ratio of the corresponding lines of two apertures, then let the expression for the intensity given by one aperture be that of Ex. 1, the expression for the other will be derived by substituting $mx$, and $my$ for $x$ and $y$, viz.

$$\left[\int m^2y\sin\left(\frac{2\pi}{\lambda}mx\sin\theta\right)dx\right]^2+\left[\int m^2y\cos\left(\frac{2\pi}{\lambda}mx\sin\theta\right)dx\right]^2.$$

Hence if $\theta_1$ makes the first a maximum or minimum, the second will be made a maximum or minimum by $\theta_2$, if $m\sin\theta_2=\sin\theta_1$, that is, if

$$\frac{\sin\theta_1}{\sin\theta_2}=\frac{m}{1}.]$$

4. The intensities transmitted to corresponding points by two similar and similarly situated apertures are in a constant ratio.

[This follows from the expression of Ex. 3, the constant ratio being $m^4$.]

5. If one aperture can be obtained from another by displacing parallel to themselves its ordinates ($y$) without altering their lengths, the intensities in the plane $xz$ are the same for the same directions.

6. The intensity at any point due to a system of equal, similar, and similarly situated apertures is equal to the intensity produced by a single aperture multi-

plied by that produced by a system of points similarly situated on the apertures, one on each.

[Consider a point situated on one of the apertures and take a point situated similarly on each of the other apertures. The vibration produced by this system of points will be of the form $A \sin(\omega t + \delta)$. As the chosen point moves over the first aperture, the corresponding points will move over the other apertures and obviously A will remain constant.

Hence if the intensity produced by a single aperture, or by a system of similarly situated points on the system of apertures, be zero, the intensity produced by the whole system will be zero. In this case then there will be in general two series of minima.]

7. Show that the pattern produced by diffraction through an elliptic aperture may be determined by reduction to the case of a circular aperture.

8. Determine the character of the diffraction pattern produced by a square aperture.

9. Determine the character of the pattern produced by diffraction through a triangular aperture.

# CHAPTER X

## ON THE POLARISATION OF LIGHT BY REFLECTION AND DOUBLE REFRACTION

**164. Transverse Vibration—Plane Polarisation.**—In the study of the phenomena with which we have been hitherto engaged, we have deduced no evidence whatsoever as to the nature of the vibrations in the luminous waves. Throughout our investigations of the phenomena of interference we have only supposed the vibrations of the interfering rays to be similar to each other, but as to how they are directed in space, or as to the relation of this direction to the direction of propagation of the ray, we have made no assumption, nor yet dealt with any phenomena calculated to attract our attention to it. From time to time our knowledge of the theory of sound has proved of considerable assistance in the study of analogous phenomena in the theory of light, but we now approach a class of optical phenomena which have no analogue in the theory of sound. These have consequently been supposed to arise from the difference in the nature of the vibrations of the ether which constitute light and those of the atmosphere which produce sound.

In the case of sound we know that the vibrations of the atmosphere are longitudinal—that is, in the direction in which the sound is being propagated; the same is the case when rods and strings vibrate longitudinally, or in the direction of their length. However, in the case of a sounding fiddle-string (or the cord of Art. 34) the vibrations of the string are perpendicular to its length, or transverse. A cord or rod is thus capable of two distinct kinds of vibrations, longitudinal and transverse, and these are propagated along the cord with different velocities. So, in general, if a disturbance is being propagated by the vibration of an elastic medium in any direction, we may resolve the vibration into two others, one longitudinal, and the other transverse to that direction, and if this applies to the ether, we should have the two species of waves in it.

The question now at issue is whether the luminiferous vibrations (supposing them to be periodic displacements) are longitudinal or transverse, or if both species exist in the ether and affect our sense of sight. This question can only be answered by experiment, but before adducing any evidence on the point, it will be well to return to the consideration of a vibrating cord (Art. 34).

Let AB (Fig. 6) represent a stretched cord, such as a fiddle-string. If this cord be rubbed in the direction of its length, it will start to vibrate in that direction, *i.e.* longitudinally, and emit a piercing note. But if a bow be drawn across it at right angles to its length, it will oscillate transversely and emit quite a different note. In the former case the appearance of the string is the same on all sides, or in all planes drawn through its length, but in the latter case the string vibrates in a definite plane, and its appearance is not the same in all planes drawn through its length. It has definite sides on it with regard to the space around it. It looks flat, like a ribbon; when vibrating it has acquired *sides* and may be said to be *polarised*.

Now suppose the string AB to pass freely through a narrow rectangular slit a little wider than the diameter of the cord. It is clear that the longitudinal vibrations will not be interfered with, but will continue unmodified, no matter how the slit is turned round the vibrating string. The case is quite different when AB is vibrating transversely, for when the length of the slit is in the direction of the vibration of the string—that is, parallel to the plane of vibration—the string is free to oscillate, and its vibrations will not be disturbed. However, as the slit is turned round the cord so as to be at right angles to the plane of vibration, the cord is no longer free to oscillate, and its motion is interrupted by the sides of the slit. Thus, by turning the slit round the cord, we detect the transverse vibrations which give a *two-sidedness* or *polarity* to the vibrating string, and even though it were invisible we should learn that it does not present the same aspect on all sides.[1]

It occurs to us now to endeavour to determine by some analogous method the nature of the luminiferous vibrations, and for this purpose we may, with profit, first operate with a plate of tourmaline cut parallel to the axis of the crystal. When a beam of light is allowed to fall perpendicularly upon such a plate part of it is transmitted, and we shall see now that this transmitted portion has the peculiar two-sided property of the transversely vibrating string. To the eye the transmitted light appears to have suffered only a slight colouring due

[1] This might also be detected by placing a smooth plane surface against the string.

to the natural tint of the crystal. But if we allow the light which passes through the plate to fall upon another similar plate, we find that the light from the first passes freely through the second when the two are parallel—that is, similarly placed with respect to the ray as A and B (Fig. 147). As the second plate is rotated (A′B′) round the beam as axis, its plane being always perpendicular to the ray, the light transmitted by it is observed to fade gradually, and when it has been turned through a right angle (A″B″) the light is completely extinguished. It is immaterial which plate is turned, but the bodily turning of both at the same time is without effect. If the rotation be continued beyond a right angle, the light reappears and increases in intensity till a second right angle has been completed, when it is as bright as at the outset. When the plates are at right angles to each other, *i.e.* crossed, the light from the first is refused transmission through the second, just as the transverse vibrations of the string were stopped by turning the slit.

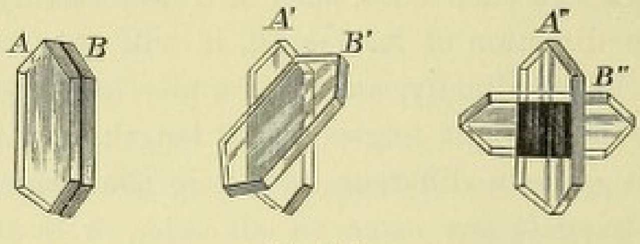


Fig. 147.

Now the beam transmitted through a single plate of tourmaline remains unaltered in intensity as the plate is rotated, we therefore detect no *two-sidedness* in ordinary or natural light; but it is quite otherwise with the transmitted beam, for we see that the amount of it which the second plate allows to pass depends on the orientation of that plate with respect to the first, and in one position it entirely refuses a passage to the light. This beam, then, which has passed through the plate of tourmaline has acquired a *two-sidedness.* It is for this reason said to be *plane-polarised*, and for this reason it has been supposed that its vibrations are transverse, and all in one plane, like the vibrations of a string plucked aside.

It is very important to remark that there is one position of the second plate which entirely cuts off the beam transmitted through the first, and this has been taken to show that the light vibrations cannot be longitudinal, for if they were, the orientation of the second plate would not be expected to affect them, just as the turning round of the slit did not affect the string vibrating longitudinally. The fact that no light gets through the second plate in one position appears to show that the longitudinal vibrations, if they exist in the ether, are not propagated as light by themselves, but it does not prove that a longitudinal component does not exist in the vibration as a necessary part, for the extinction of the transverse constituent might entail

Longitudinal component.

that of the longitudinal as well. In fact, it is not easy to see how waves of transverse vibration can be propagated in a medium such as an elastic solid without being attended by more or less of the longitudinal element. Further evidence on this point will appear in what follows.

**165. Polarisation by Reflection—Biot's Polariscope.**—The polarisation of light was first noticed by Huygens when studying the refraction of light through a crystal of Iceland spar, but it remained an isolated fact in science for more than a century afterwards. About 1808 Malus discovered, accidentally, that light when reflected from the surface of glass acquires properties similar to those possessed by the light transmitted through a plate of tourmaline—that, in fact, light may be polarised by reflection,—and pursuing the inquiry further, he found that the same occurs when light is reflected from water and other transparent substances. Hence the two-sidedness of the light which has passed through a tourmaline plate may be detected by allowing it to fall upon a plane glass plate. By turning the plate round the beam the reflected light is seen to vary in intensity, and in one position of the plate the reflected light vanishes altogether. By keeping the glass plate stationary and rotating the tourmaline we may obtain the same results.

Similarly the beam may fall first upon the glass and afterwards be transmitted through the tourmaline with the same effect. In one case we *polarise* the light by transmission through the tourmaline and *analyse* it (that is, detect its polarisation) by reflection from the mirror, and in the other case it is polarised by reflection from the mirror and analysed by transmission through the tourmaline. It is clear, therefore, that the tourmaline may be dispensed with altogether and replaced by a plane glass mirror, since the mirror can act the part either of a polariser or an analyser. On this principle instruments termed *polariscopes* have been constructed. One of the first of these was designed by Biot,[1] but it has long since been superseded by more commodious forms. It is represented in Fig. 148. At each end of a tube T plane mirrors of polished black glass are placed. Each mirror is capable of two motions—one round a diameter of the tube—that is, round an axis perpendicular to the axis of the tube. The amount of this rotation is

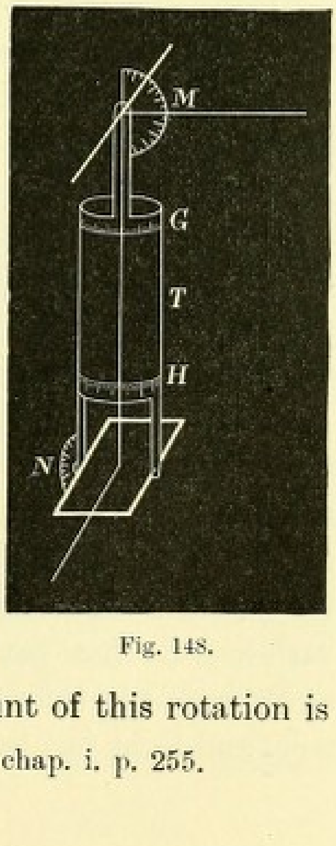


Fig. 148.

[1] Biot, *Traité de Physique*, tom. iv. livre sixième, chap. i. p. 255.

measured by the graduated circles M and N. The second motion is round the axis of the tube. This is obtained by the mirrors being attached to rings G, H, which are graduated and movable on the tube.

The mirrors can thus be inclined at any angle to the incident light and to each other.

**166. Angle of Polarisation.**—When the planes of the two mirrors are placed parallel, the light which is reflected from the first and falls upon the second is, for a certain incidence, entirely reflected there. By rotating either mirror round the axis of the tube, the amount of reflected light will diminish and become a minimum when the mirror has been rotated through 90°, so that the mirrors are crossed. The amount of light reflected in this position will depend upon the angle of incidence, and for one particular value of this angle the reflected light will vanish entirely.

We consequently infer that there is one particular angle of incidence at which light is completely polarised by reflection from glass, and this is termed the *angle of polarisation*, or the *polarising angle*.

When light is reflected from glass, the reflected beam in general is
rtial. only partially polarised. It cannot be completely extinguished by a tourmaline plate or by reflection from another mirror at any incidence. The amount of polarisation depends upon the angle of incidence, and at one particular angle the polarisation becomes complete.

We must not hastily infer that for every substance there is an angle of complete polarisation. In fact, it is proved by experiment that as the angle of incidence increases, the polarisation in general also
ximum. increases to a maximum, and then decreases after passing through the angle of maximum polarisation. For each substance there is an angle of incidence which gives a maximum of polarisation, and this angle is termed the *polarising angle* of the substance.

M. Jamin, who investigated this subject, found that only a few substances, of refractive index about 1·46, polarise light completely by reflection. For all other substances the polarising angle is merely the angle of maximum polarisation. For glass the polarising angle is about 57°, and for pure water 53° 11′.

**167. Plane of Polarisation.**—When plane-polarised light falls at the polarising angle upon a glass mirror, the intensity of the reflected light depends upon the position of the *plane* of incidence with regard to the ray. Thus as the mirror is rotated round the ray, keeping the angle of incidence constant the intensity of the reflected pencil varies, and that particular plane of incidence in which the light is most copiously reflected is called *the plane of polarisation*. When the polarised light has been obtained by refraction, it is obvious that the plane of polarisation, as defined above, is the plane of reflection of the light, for this beam would be most copiously reflected by a second surface when it is parallel to the polarising surface.

According to the theory of Fresnel, the vibrations of plane-polarised light are perpendicular to the plane of polarisation. Thus the direction of vibration in light, polarised by reflection, is perpendicular to the plane of reflection—that is, parallel to the reflecting surface. The relation of the direction of vibration to the plane of polarisation has been a subject of much dispute, and we postpone its consideration for the present. It will be well to bear in mind, however, that something *may be* going on both in and perpendicular to the plane of polarisation, and that the vibration may not be a simple displacement in the wave front (chap. xxi.). Direction of vibration.

**168. Double Refraction.**—Hitherto we have assumed that when light is incident on the surface of a transparent medium, the refracted portion pursues in all cases a single definite direction. That this is not always the case was discovered by Erasmus Bartholinus, a Danish philosopher, about the year 1669. Experimenting with a crystal of Iceland spar (carbonate of calcium), he found that a beam of light on refraction at its surface travels through the crystal in two determinate pencils, one of which traverses it according to the known laws of refraction, while the other is bent according to a new and extraordinary law not hitherto noticed.[1]

A few years after the discovery of double refraction, Huygens[2] gave a satisfactory explanation of it in uniaxal crystals on the principles of the wave theory, and whilst repeating the observations of Bartholinus, he was led to the discovery of the "wonderful phenomenon" of the polarisation of light.

The property of producing two refracted beams is called *double refraction*, and is possessed by all crystallised minerals except those whose fundamental form is the cube. It belongs also to animal and vegetable substances possessing a regular arrangement of parts, and to all bodies whose parts are in a state of unequal compression or dilatation.

The angular separation of the two refracted pencils varies with the direction of the incident ray with reference to the natural figure of the crystal. In every doubly refracting crystal there is at least one direction, and in many two, in which no such separation occurs. These Optic axi

[1] An account of these experiments was published at Copenhagen in 1669 under the title *Experimenta Crystalli Islandici dis-diaclastici, quibus mira et insolita refractio detegitur*. A full account of them is given in the *Edinburgh Philosophical Journal*, vol. i. p. 271.

[2] Huygens, *Traité de la Lumière*, "De l'estrange refraction du Cristal d'Islande."

directions are called *optic axes* of the crystal. The refracted rays are most widely separated when the incident beam is perpendicular to the optic axis.

In Iceland spar, the substance in which it was first observed, the phenomenon of double refraction is very strikingly exhibited. This mineral crystallises in many forms, all of which may be reduced by cleavage to the rhombohedron, which is accordingly the fundamental form. It is also found in considerable masses of great purity and transparency. The rhombohedron is bounded by six parallelograms, the angles of which are 101° 55′ and 78° 5′ respectively. Two of the solid angles, *a* and *b* (Fig. 149), diametrically opposite, are contained by three obtuse angles, while each of the remaining four is bounded by one obtuse and two acute angles. The direction making equal angles with the faces at the summits of the obtuse solid angles is called the *axis of the crystal*. Thus if the edges of the rhomb be equal to each other, the line *ab* adjoining the obtuse solid angles, or any parallel line, is the crystallographic axis. The angles at which the faces themselves are inclined are 105° 5′ and 74° 55′.

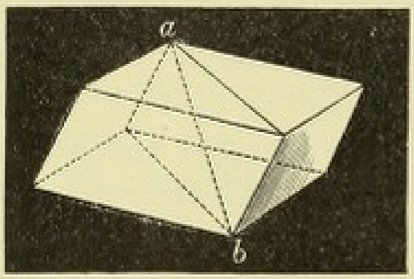


Fig. 149.

When a transparent crystal of Iceland spar is placed over a small black dot on a sheet of white paper, two images of the dot are seen on looking through the crystal, and if the eye be held perpendicularly over the face of the crystal while it is rotated over the dot, one of the images remains stationary, while the other rotates round it. The fixed image appears a little nearer than the movable one, and the line joining them is parallel to the shorter diagonal of the rhombic face through which they are observed.

**169. Polarisation by Double Refraction.**—The properties of light, which has suffered double refraction, may be best examined by allowing a pencil from a small aperture to pass through a large crystal so as to receive the two emerging beams on a screen.

Of the two portions into which the refracted light is divided by a uniaxal crystal, that which obeys the ordinary laws of refraction is called the *ordinary ray*, and this gives an image on the screen called the ordinary image. The other refracted ray does not obey the ordinary laws of refraction. It is called the *extraordinary ray* and gives an extraordinary image. These rays and images may be denoted by the letters O and E. By placing a plate of tourmaline in the path of either of the rays produced by double refraction, it is found on rotating the tourmaline that in one position it refuses to transmit

oth eams ane- larised.

the ray, while in other positions the ray is transmitted with more or less freedom. Both the rays then are two-sided or plane-polarised.

This may also be shown by allowing the pencils to pass through a second crystal of calcspar. When the second crystal is parallel to the first, the two behave just like one of their joint thickness. In this position the ordinary and extraordinary rays of the first crystal traverse the second as ordinary and extraordinary rays, but when the second is rotated through an angle, each of the rays O and E is doubly refracted in it, so that four rays emerge from the second crystal and four images are depicted on the screen. The ordinary ray of the first crystal gives rise to an ordinary $O_o$ and an extraordinary $O_e$, while the extraordinary gives rise to an ordinary $E_o$ and an extraordinary $E_e$. As the rhomb is turned, the pair $O_e$ and $E_o$, which are faint at first, gradually increase in brightness, while the other pair $O_o$ and $E_e$ diminish in intensity and finally vanish when the crystal has been rotated through 90°. We have now again only two images, viz. $O_e$ and $E_o$. On rotating the rhomb farther, these images grow fainter, while the other pair $O_o$ and $E_e$ reappear and increase in intensity till the rhomb is rotated through 180° from its first position. Here we have only two images $O_o$ and $E_e$ as at starting. Thus in one position the second rhomb allows the pencils to pass freely through—the ordinary as ordinary, and the extraordinary as extraordinary,—while in a position at right angles to this it refracts the ordinary ray entirely as an extraordinary, and the extraordinary entirely as an ordinary and in any intermediate position it refracts both doubly.

By utilising the persistence of visual impressions, this form of the experiment is susceptible of a very elegant modification. Let one of the pencils emerging from the first crystal fall upon the second. This beam will in general be divided into two others by the second, and the refracted pencils will depict two spots on a screen placed to receive them. By placing the screen at a proper distance the spots may be made to partly overlap. Now let the second crystal be rotated rapidly around the direction of the incident pencil of light. The spots on the screen will describe circular luminous bands on the screen which partly overlap. Three concentric rings are consequently presented. The middle ring is due to the overlapping of the two images, and is uniformly bright all round, for the overlapping images are complementary in all positions. This central ring is bordered on either side by other rings, one due to the ordinary image and the other to the extraordinary. These do not appear uniformly bright. The former will be brightest in the plane of polarisation, and darkest in the perpendicular plane. The reverse is the case in the extraordinary ring.

The ordinary and extraordinary rays emerging from a rhomb of Iceland spar or any other doubly refracting crystal are consequently in a condition singularly different from that of common light, for while a beam of natural light is always divided into two of equal intensity on entering the crystal (except when it passes in one particular direction called the optic axis), the subdivision of the ordinary or extraordinary ray by a second rhomb depends on the orientation of the second crystal with respect to the first.

It was in this form that the polarisation of light was first noticed by Huygens. On analysing his observations, he determined that a beam of solar light is always divided into two (except when it traverses the crystal in a direction called the optic axis), and that each of the resulting beams will be singly or doubly refracted by a second crystal of spar according to the relative position of the *principal sections;* these are planes drawn through the optic axes of the crystals and perpendicular to the refracting faces. If the principal planes of two crystals be either parallel or at right angles to each other, then the rays which emerge from the first are not doubly refracted by transmission through the second. If the principal planes are parallel, the ordinary ray from the first traverses the second as an ordinary ray, and the extraordinary as extraordinary; but if these sections are at right angles, the ordinary ray from the first is refracted as an extraordinary ray in the second, and the extraordinary ray as an ordinary.

The ordinary wave is found to be polarised in the principal plane, and the extraordinary wave is polarised perpendicularly to the principal plane, the plane of polarisation being defined as in Art. 167.

A beam of plane-polarised light possesses the following characteristics:—

I. It is not divided into two others by a doubly refracting crystal in two positions of the principal section with respect to the ray, while in other positions it is divided into two pencils which vary in intensity, and are complementary as the crystal is rotated.

II. It is not reflected at the surface of a transparent substance when the plane of incidence is perpendicular to the plane of polarisation, and when the angle of incidence has a certain value depending on the nature of the substance.

**170. Property of Tourmaline.**—In our preliminary experiment we obtained polarised light by transmitting a pencil of ordinary light through a plate of tourmaline. Now tourmaline is really a doubly refracting crystal and divides the intromitted beam into two parts, an ordinary and an extraordinary, yet it is only the extraordinary beam that emerges from the plate, for the ordinary pencil is rapidly absorbed by the crystal, so that a plate of small thickness (1 or 2 mm.) is almost impervious to it. It is this property that renders tourmaline plates so readily adapted to the operations of either producing polarised light or of analysing it when already obtained.

The wave velocity is then not the only property affected by crystalline structure. In many crystals the two polarised rays suffer different rates of absorption, and it is this property that qualifies a single plate of tourmaline to act the part either of a polariser or an analyser. The property of double absorption is without doubt intimately related to that of double refraction. Double absorptio

Tourmaline, however, is not a very transparent mineral, and strong beams of polarised light cannot well be obtained with it. Hence it is quite unfitted for work when the illumination is faint, consequently other forms of polarisers and analysers—that is, apparatus for producing and studying polarised light—have been invented. The most important of these are Nicol's and Foucault's prisms (see Arts. 182, 183).

**171. Brewster's Law.**—About the year 1811 Sir David Brewster commenced an extensive series of experiments with the object of determining the polarising angles of different media, and of connecting them by a law. The outcome of his investigations was the simple and remarkable law, "*The index of refraction of the substance is the tangent of the polarising angle.*" This law, which is expressed by the formula

$$\tan i = \mu,$$

informs us that the polarising angle ranges in different substances from 45° upwards, and, when the refractive index is known, the angle of polarisation is inferred.

Since the refractive index is different for differently coloured lights, it follows that the angle of maximum polarisation is different for the different rays of the spectrum, and consequently if a beam of *solar* light be reflected successively from two glass plates whose planes of reflection are at right angles, the reflected beam will never be wholly extinguished, but will be reduced to a residuum coloured with a red or blue tinge, according as the angle of incidence is the polarising angle of the more or less refrangible rays. When the angle of incidence is that of polarisation of the most luminous part of the spectrum, the reflected light is of a purplish tint, formed by the mixture of the remaining rays in different proportions. These effects are best marked in highly dispersive substances. Colour at polarising angle.

The geometrical interpretation of Brewster's law is that when a pencil of light falls upon a transparent substance at the polarising angle, *the reflected and refracted rays are at right angles.* For if

$$\tan i = \mu, \quad \text{then} \frac{\sin i}{\cos i} = \frac{\sin i}{\sin r},$$

or

$$\cos i = \sin r \quad \text{and } i + r = 90^\circ,$$

therefore $i$ and $r$ are complementary, or the reflected and refracted rays are at right angles.

**172. Pile of Plates.**—The law of Brewster applies to light reflected from the surface of the rarer as well as from the surface of the denser medium, and since the refractive index in the former case is the reciprocal of that in the latter, it follows that the angles of polarisation in the two cases are complementary. From this it follows that when a beam of ordinary light falls upon a parallel plate of any transparent substance at the polarising angle, the refracted portion will meet the second surface at the polarising angle, and if it still contains an unpolarised part this will be wholly or partially polarised by reflection there. Hence if several plates of glass be arranged parallel to one another, a pencil of light incident on the first at the polarising angle will, after refraction, meet all the succeeding surfaces at their polarising angles also. So that all the light reflected from these surfaces will be plane-polarised. Such an arrangement is termed a *pile of plates*, and is very useful as a polariser when the light is not incident in a parallel beam as in the case of skylight, for the reflected beam is much more intense than that obtained from a single plate.

**173. Polarisation of the Refracted Light.**—So far we have only considered the modification which the reflected pencil has suffered. However, when the refracted pencil is examined, it is found to contain a quantity of polarised light. The relation between this polarised light and that of the reflected beam is very intimate. It was discovered by Arago and stated thus: "*When an unpolarised ray is partly reflected at, and partly transmitted through, a transparent surface, the reflected and transmitted portions contain equal quantities of polarised light, and their planes of polarisation are at right angles to each other.*"

The operation of the plate is purely selective, for the polarised component, which is missing in the reflected light, is represented in the transmitted light.

If the transmitted light be received upon a second parallel plate, the portion of common light which it contains undergoes a further subdivision, and so on for any number of plates. Hence, when the number of plates is sufficiently great, the transmitted light will be completely polarised, and consequently if a beam of light be incident on a pile of plates the transmitted light, after traversing a certain number of the plates, will suffer no further diminution in intensity except by actual absorption in the plates. For the refracted light will become wholly polarised in a plane perpendicular to the plane of incidence, and no portion of it will be reflected by the succeeding plates.

**174. Law of Malus.**—These results were established by Malus, who also conjecturally assumed that when a pencil, polarised by reflection at one plane surface, is allowed to fall upon a second at the polarising angle, *the intensity of the twice-reflected beam varies as the square of the cosine of the angle between the two planes of reflection.* The truth of this law was afterwards established by the experiments of Arago.

Thus if we assume with Fresnel that the direction of vibration is perpendicular to the plane of polarisation, then if the incident light be polarised in a plane making an angle $\theta$ with the plane of incidence, the incident vibration, of amplitude $a$, may be resolved into two components, one $a \cos \theta$ perpendicular to the plane of incidence and the other $a \sin \theta$ parallel to it. The former is polarised in the plane of incidence, and is reflected; the latter is polarised at right angles to the plane of incidence, and is transmitted. The reflected light is thus in all cases polarised in the plane of reflection, and its intensity is proportional to $\cos^2 \theta$. The law of Malus is thus simply accounted for by the principle of resolution of vibrations.

*Ordinary light*, we know, is broken up by a crystal into two parts polarised at right angles to each other, and conversely, the preceding law of Malus enables us to regard ordinary light as consisting of two equal plane-polarised rays polarised in planes at right angles. For consider two such rays, and let $\alpha$ and $90° - \alpha$ denote the angles which their planes of polarisation make with the plane of reflection; then if I denotes the intensity of each of the rays, the intensities of the reflected rays are $\mathrm{I} \cos^2 \alpha$ and $\mathrm{I} \sin^2 \alpha$ respectively, and the sum of these,

$$\mathrm{I} \cos^2 \alpha + \mathrm{I} \sin^2 \alpha = \mathrm{I},$$

is the resultant intensity of the reflected. It is therefore constant and independent of the position of the plane of reflection with respect to the ray; but this is the distinctive characteristic of ordinary unpolarised light.

Hence in ordinary light the direction of vibration is supposed to be quite irregular, but when it falls upon a doubly refracting medium the vibrations seem to be resolved in two definite directions, constituting two equally intense rays polarised in perpendicular planes and differently refracted by the medium.

The foregoing law of Malus is equivalent to stating that if a plane-polarised ray be reflected at a surface, it is only the component of the vibration parallel to the surface that is reflected. The component parallel to the plane of incidence cuts down into the surface, as it were, and is refracted.

If a pencil of light of intensity I falls upon a crystal of calcspar, it is divided into ordinary and extraordinary rays, each of intensity $\frac{1}{2}$I, and if these fall upon a second crystal, they will be again subdivided according to the law of Malus as indicated in the following table:—

| Incident light. | In first crystal. | In second crystal. | Plane of polarisation. | Sum. |
|---|---|---|---|---|
| I | $O = \frac{1}{2}I$ | $O_o = \frac{1}{2}I\cos^2\alpha$ | 0 | $\frac{1}{2}I$ |
| | | $O_e = \frac{1}{2}I\sin^2\alpha$ | 90 | |
| | $E = \frac{1}{2}I$ | $E_o = \frac{1}{2}I\sin^2\alpha$ | 0 | $\frac{1}{2}I$ |
| | | $E_e = \frac{1}{2}I\cos^2\alpha$ | 90 | I |

**175. Partially Polarised Light.**—When light is incident on a reflecting surface at an angle either greater or less than the polarising angle, it was observed by Malus that the reflected light possesses only in part the properties of polarised light. Neither of the two pencils into which it is divided by a rhomb of Iceland spar ever completely vanishes, but each varies in intensity as the rhomb is rotated. He consequently concluded that the reflected light consisted of two parts, one perfectly polarised, while the residue remains in the state of ordinary light. Partially polarised light is then, according to Malus, a mixture of ordinary light with a part wholly polarised, and in this hypothesis he has been followed by most subsequent philosophers, for the light possesses all the properties of such a mixture.

If this partially polarised light be reflected from a second surface at the same angle, the reflected pencil is found to contain an increased quantity of polarised light, and by augmenting the number of reflections the light may be almost wholly polarised. This was first noticed by Sir David Brewster, and he found that light may be polarised at any incidence by a sufficient number of reflections, the number of reflections required increasing as the angle of incidence is more removed from the polarising angle. Hence the utility of a pile of plates.

**176. Interference of Polarised Light.**—We have seen that light suffers an important modification by transmission through doubly refracting crystals and also by reflection, and it has also been found that this modified or polarised light, as it has been called, obeys the ordinary laws of reflection, refraction and dispersion. It is important therefore to determine if the phenomena of interference are produced by it under the same circumstances as by ordinary light. For this purpose the refractometer of M. Billet, already described (Art. 103), will be found convenient. A beam of plane-polarised light produced by a tourmaline T, or otherwise, falls upon a narrow slit S (Fig. 150). The light diverging from the slit falls upon the segments L and L′ of a divided lens, which bring the rays to foci A and B. At these points thin plates of Iceland spar, or tourmaline, or other doubly refracting crystals may be placed. If tourmalines be at A and B there is a single polarised beam from each, and it is found that when their principal planes are parallel we have interference fringes on the screen, as in the case of ordinary light, but when one of them is rotated so that they are crossed the fringes disappear entirely. This is what we should have expected, for we have already learned (Art. 47) that two rectangular vibrations differing in phase in general compound into an elliptic vibration.

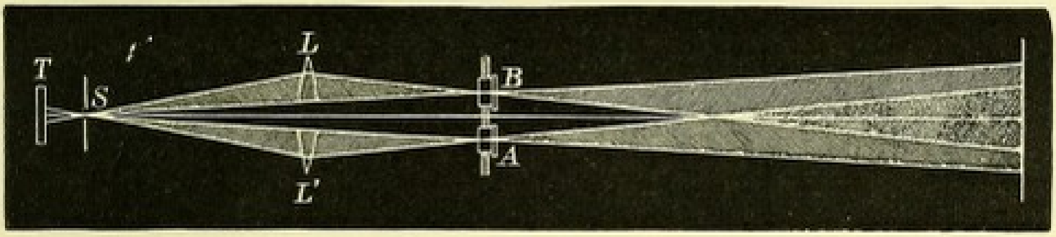



Fig. 150.

If thin plates of Iceland spar be placed at A and B with their principal sections parallel, and if the incident light be polarised in or perpendicular to their principal sections, it will traverse the crystals without double refraction, either as ordinary or as extraordinary rays. The planes of polarisation of the emerging beams will be parallel and interference fringes will be formed on the screen. If however the incident light is polarised in any other plane, double refraction will occur in the plates at A and B, each giving rise to an ordinary and to an extraordinary pencil. The two ordinary beams interfere and produce a system of fringes. Superposed on these fringes we have another system produced by the extraordinary beams, but no destructive interference occurs between the ordinary and extraordinary parts.

If now one of the plates A and B be turned through 90°, so that they are crossed, the ordinary ray from A will be polarised parallel to the extraordinary ray from B, and these will interfere and produce a system of fringes. The centre of the system will however be displaced towards A, for the ordinary ray travels more slowly in the crystal than the extraordinary, as we shall see immediately. Similarly the extraordinary ray from A will interfere with the ordinary from B and produce a system displaced towards B.

Conditions for interference.

Fresnel and Arago, who investigated directly the interference of polarised light, summarised their conclusions as follows.[1]

(1) Two rays of light polarised at right angles do not interfere destructively under the same circumstances as two rays of ordinary light.

(2) Two rays of light polarised in the same plane interfere like two rays of ordinary light.

(3) Two rays polarised at right angles may be brought to the same plane of polarisation without thereby acquiring the quality of being able to interfere with each other.

(4) Two rays polarised at right angles, and afterwards brought to the same plane of polarisation, interfere like ordinary light if they originally belonged to the same beam of polarised light.

The fact that rays polarised in perpendicular planes cannot interfere destructively is in itself an indication that the direction of vibration is transverse to the direction of propagation.

[1] Fresnel, *Œuvres*, tom. i. p. 521.

## CHAPTER XI

### DOUBLE REFRACTION IN UNIAXAL CRYSTALS

**177. Wave Surface in Uniaxal Crystals.**—Before proceeding to the general theory of double refraction and the more complicated phenomena arising from refraction in biaxal crystals, it will be well to first give a general statement of the phenomena and laws of refraction in uniaxal crystals, and it will be interesting afterwards to see how these may be deduced as particular cases of the general theory when we suppose the two optic axes of a biaxal crystal to coincide.

It has been already stated that double refraction was discovered in Iceland spar by Erasmus Bartholinus. Soon after its discovery Huygens,[1] who had already unfolded the wave theory of light and accounted for ordinary refraction and reflection, was naturally anxious to reconcile the new properties of light discovered by Bartholinus with the same theory, and in his desire to assimilate the two classes of refraction he was happily led to assign the true law of *extraordinary refraction* in uniaxal crystals. He had already shown that the form of the wave of light propagated in glass and isotropic substances was a sphere, and as one of the rays in Iceland spar was found to obey the ordinary laws of refraction he assumed that the corresponding wave was also a sphere. The law which governed the other ray, though not so simple, he imagined to be next in order of simplicity, and he assumed the extraordinary wave to be a spheroid—that is, an ellipsoid of revolution.

The velocity of the extraordinary ray in any direction is consequently given by the following construction:—"Let an ellipsoid of revolution be described round the optic axis having its centre at the point of incidence; and let the greater axis of the generating ellipse be to the lesser in the ratio of the greatest to the least index of refraction: then the velocity of any ray will be represented by the radius vector of the ellipsoid which coincides with it in direction."

[1] Huygens, *Traité de la Lumière*, chap. v.

This law was found to apply to many crystals besides Iceland spar, but in all of these there was only one optic axis, or one direction along which a ray of light passed without division. The researches of Brewster, however, made known a class of crystals having two optic axes or two directions of no separation of the ray. Huygens's law was then found not to be general, and in this state the problem was taken up by Fresnel, who proposed a theory which met not only all the requirements of the ascertained facts of the refraction in biaxal crystals but which even outran the existing knowledge and predicted results of the highest consequence, afterwards verified by direct observation.

Two sheets. According to the construction of Huygens the wave surface or the secondary wave in a uniaxal crystal consists of two portions or sheets, one a sphere which gives rise to the ordinary ray, and the other a

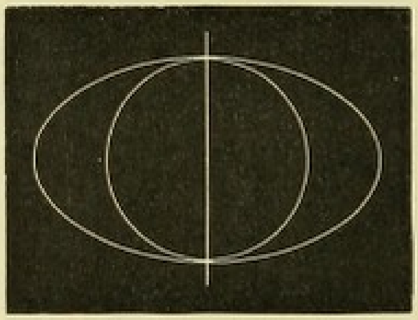


Fig. 151.—Negative Crystals.

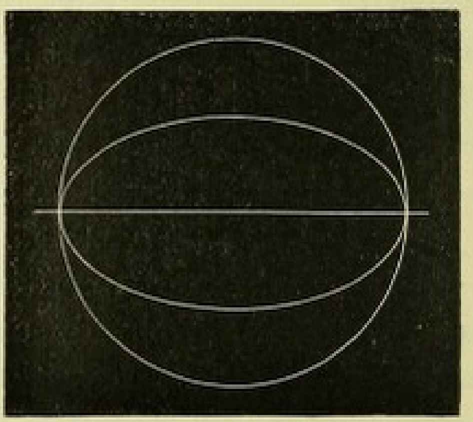


Fig. 152.—Positive Crystals.

spheroid giving rise to the extraordinary ray. These two surfaces, the sphere and spheroid, touch at two points and the line joining these points is the optic axis.

In the case of Iceland spar and all *negative* crystals the sphere is entirely within the spheroid, and in the case of quartz and *positive* crystals the spheroid is within the sphere. Thus if the spheroid is generated by the revolution of the ellipse

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}=1$$

round its minor axis $b$, and the sphere by the revolution of the circle

$$x^2+y^2=b^2$$

round the same axis, any section of the wave surface through the optic axis ($b$) is an ellipse (axes $a$ and $b$) and a concentric circle of radius $b$ touching it at the extremities of the axis minor (Fig. 151).

In positive crystals the wave surface is generated by the revolution of the ellipse round its axis major and a concentric circle of radius $a$ touching it at the extremities of that axis (Fig. 152).

A section of the complete wave surface by a plane perpendicular to the optic axis consists of two circles, one of radius $a$ and the other of radius $b$.

Huygens verified his theory by well-contrived experiments, but much more accurate measurements were necessary to prove the extraordinary wave to be truly an ellipsoid of revolution. These measurements were made in 1802 by Wollaston, and afterwards by Stokes, Mascart, and Glazebrook with the most perfect optical instruments, resulting in the complete verification of Huygens's hypothesis.

**178. Huygens's Construction.**—Let us now seek the directions of the two refracted rays in a crystal of calcspar when a plane wave falls upon it from air. Let IA (Fig. 153) be the direction of the light incident on the face of the crystal, and let AB be the trace on the plane of the paper of the incident wave front, and AA′ the trace of the face of the crystal. The incident wave front is a plane through AB perpendicular to the paper, and the face of the crystal a plane through AA′, also perpendicular to the plane of the paper.

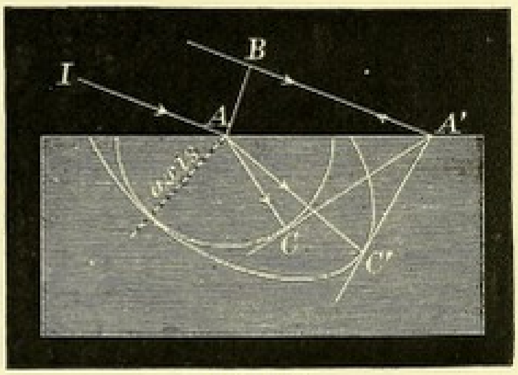


Fig. 153.

When the disturbance reaches A this point becomes the centre of a spherical wave reflected back into the air, and also the centre of a double wave propagated in the crystal. The plane of the paper will cut this wave in two curves, the sphere in a circle and the spheroid generally in an ellipse. The diagram represents the particular case in which the optic axis lies in the plane of the paper, and consequently the circle and ellipse touch each other. If the disturbance from B reaches A′ at the instant the wave in the crystal is just developed to the extent represented in Fig. 153, then through a perpendicular drawn at A′ to the plane of the paper—that is, through the line in which the incident wave meets the surface—draw a tangent plane to the sphere. This plane will be the ordinary wave front. It will touch the sphere at C in the plane of the paper, and AC will be the ordinary refracted ray. Through the same line draw a tangent plane to the spheroid. This plane will be the front of the extraordinary wave, for all the wave surfaces diverging from the various points of AA′ at the same instant will touch it. If this plane touches the spheroid at a point C′, then C′ will not in general lie in the plane of the paper, unless the optic axis lies in (or is perpendicular to) that plane; AC′ will be the extraordinary ray, and in general it will not be in the plane of incidence or obey the ordinary laws of refraction. If the optic axis lies in the plane of the paper—that is, the plane of incidence, as in the diagram—the point C′ will also lie in that plane, so that the extraordinary ray AC′ will lie in the plane of incidence and will thus obey one of the laws of refraction.

There is one case, however, in which the extraordinary ray should obey both laws of refraction, viz. when the optic axis is perpendicular to the plane of incidence. In this case the section of the spheroid by the plane of incidence is equatorial, and is therefore a circle, so that the extraordinary ray AC′ is not only in the plane of the paper, but the sine of the angle of incidence bears a constant ratio to the sine of the angle of refraction. This ratio is called the *extraordinary index* of refraction. If the velocity in air be denoted by unity, and the velocities of the ordinary and extraordinary rays by $b$ and $a$, the radii of these circles, then

$$\mu_o = 1/b = \text{ordinary index,}$$
$$\mu_e = 1/a = \text{extraordinary index.}$$

The extraordinary index of refraction is thus the least value of the index of refraction of the extraordinary ray in negative crystals.

**179. Verification of Huygens's Construction.**—The simplest and most exact method of showing that one of the rays in Iceland spar obeys the ordinary laws of refraction, no matter in what direction it traverses the crystal, and consequently that its wave is spherical, is to cut several slices in different directions from a rhomb of spar and to cement them together and then to cut the whole into a prism, having its edges perpendicular to the planes of junction (Fig. 154). Examining through this prism the light from the slit of a spectroscope, the extraordinary spectra furnished by the different slices are observed to be differently deviated, but there is only one ordinary spectrum, or the ordinary spectra furnished by the different slices all coincide—that is, the refractive index of the ordinary ray is independent of the direction in which it traverses the crystal. Its wave surface is therefore spherical; and its refractive index $\mu_o$ is measured in the same manner as that of any uncrystallised substance.

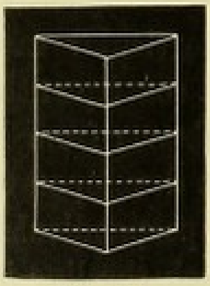


Fig. 154.

To verify the construction for the extraordinary wave, we examine the following cases.

(1) *Refracting face parallel to the optic axis, and the plane of incidence perpendicular to the axis.*—Let the face of the crystal be a plane

through AA′ (Fig. 155) perpendicular to the plane of the paper (which is supposed to be the plane of incidence). Then in this case the optic axis at A is a line through A perpendicular to the plane of the diagram. The section of the sphere is a circle of radius $b$, and the section of the spheroid, being its equatorial section, is a circle of radius $a$. The tangent planes from a line through A′ perpendicular to the plane of the paper, to the sphere and spheroid, touch them at C and C′ in the plane of incidence. Taking the velocity of light in air as unity, the velocity of the ordinary ray will be $b$, the velocity of the extraordinary in this case will be $a$, and the refractive index for the extraordinary ray is

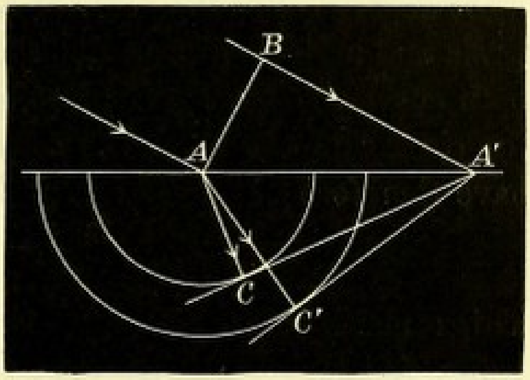



Fig. 155.

$$\frac{\sin i}{\sin r} = \frac{1}{a} = \mu_e.$$

It therefore obeys both laws of refraction.

Cutting a prism of Iceland spar with its refracting edge parallel to the optic axis, two spectra are obtained; the light of one is polarised in the principal plane and the other in the perpendicular plane. By interposing a plate of tourmaline in the path of either it can be cut off and the other examined. The indices $\mu_o$ and $\mu_e$ can thus be calculated for the several rays of the spectrum. The results of experiment are in complete accordance with the foregoing theory, and the section of the extraordinary wave perpendicular to the optic axis is therefore a circle of radius $a = 1/\mu_e$.

This wave is then a surface of revolution round the optic axis. To determine the form of the generating curve, we shall study the refraction in a plane passing through the optic axis.

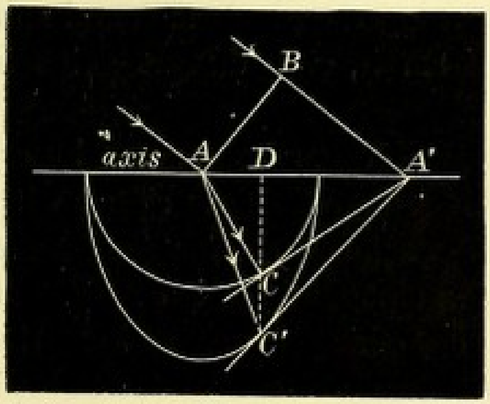



Fig. 156.

(2) *Optic axis parallel to the face of the crystal and to the plane of incidence.*—When the refracting surface contains the axis of the crystal and the plane of incidence passes through that axis, the section of the spheroid by the plane of incidence will be an ellipse whose lesser axis (the optic axis) lies in the surface (Fig. 156). The section of the sphere will be a circle touching the ellipse at the extremities of the minor axis. The

tangent planes from A′, as before, will touch the sphere and spheroid in C and C′ respectively in the plane of incidence, and since A′ is on the axis minor it follows that the line CC′ will meet AA′ at right angles. For the polar of any point on the chord of contact of a circle and ellipse having double contact is the same with regard to both curves. Hence CC′ must be the polar of A′ with regard to both the circle and ellipse, and is consequently perpendicular to AA′. We have therefore

$$\frac{\tan r}{\tan r'} = \frac{DC'}{DC} = \frac{a}{b} = \frac{\mu_o}{\mu_e}.$$

This remarkable relation has been verified by Malus[1] as follows:—Two scales AC and BC (Fig. 157) engraved on a plate of polished steel are inclined to each other at a small angle and divided into small

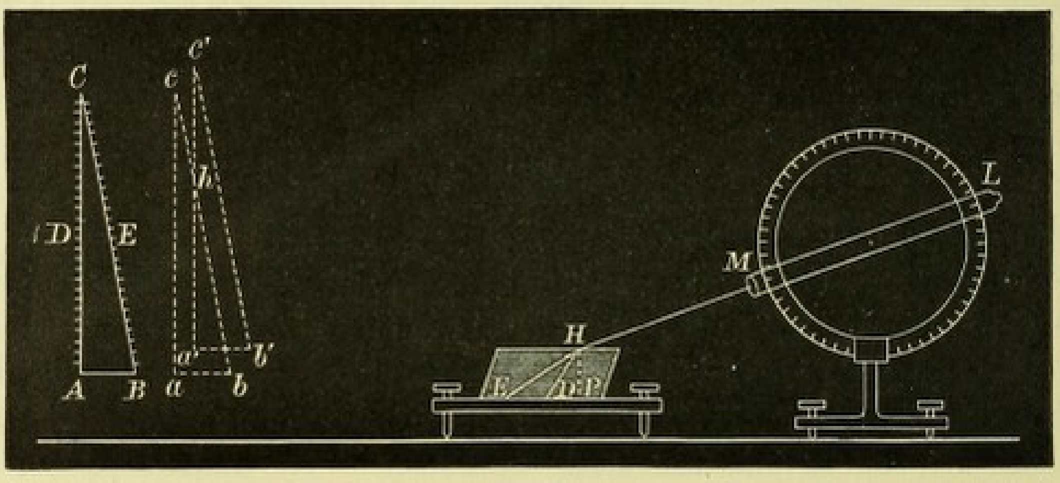



Fig. 157.—Verification of Huygens's Wave Surface.

equal parts. A thick plate of the crystal having its faces parallel to the optic axis is laid on the scale and viewed through a telescope LM mounted on a graduated vertical circle. The scale and crystal are supported on a horizontal circle, and the horizontality of the upper face of the crystal is ensured by turning the platform round and observing that the image, by reflection from it, of a distant point is not displaced.

Two images of each scale are seen, and if these be denoted by *ac*, *a′c′*, *bc*, *b′c′*, there will in general be some point of *bc* coinciding with some point of *a′c′*. Let this point be *h*. Then *h* is the image of some point D of the scale AC, and also of some point E of the scale BC. If the axis of the telescope is directed to view this point, it will cut the surface of the crystal at H and the position of H can be determined with reference to the scales. The divisions at E and D which appear to coincide can be read off and the distance ED determined by actual measurement.

[1] "Théorie de la double réfraction" (*Mémoires des Savants étrangers*, tom. ii. p. 303).

If $e$ be the thickness of the crystal HP we have

$$ED = EP - DP = e (\tan r' - \tan r).$$

But $\tan r$ is known, for the angle of incidence is equal to the inclination of HM to the vertical, and is consequently equal to the angle made with the vertical by the axis of the telescope, and $\sin i = \mu_o \sin r$; therefore $r$ is known and $r'$ may be determined by the above formula. If this value of $r'$ agrees with its value as determined by the formula,

$$\frac{\tan r'}{\tan r} = \frac{\mu_e}{\mu_o},$$

the experiment will have proved that the section of the extraordinary wave by a plane through the optic axis is an ellipse, and consequently, since the wave front is a surface of revolution, it must be a spheroid of axes $a$ and $b$. The experimental results are here in exact accordance with the theory.

(3) *Optic axis perpendicular to the refracting surface.*—When the optic axis is perpendicular to the face of the crystal it must be parallel to the plane of incidence, and the section of the wave surface should be a circle of radius $b$ and an ellipse of axes $a$ and $b$, as in Fig. 158. If a circle of radius $a$ be described with centre A it will touch the ellipse at the extremities PQ of its axis. A tangent from A′ to this circle will touch it at C″, and if the tangent to the ellipse touch it at C′ then C′C″ will be perpendicular to the axis PQ of the ellipse. Consequently if the angle OAC″ (= AA′C″) be denoted by $\rho$, we have

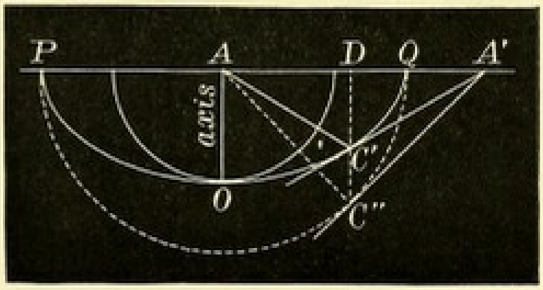



Fig. 158.

$$\frac{\tan \rho}{\tan r'} = \frac{DC'}{DC''} = \frac{b}{a} = \frac{\mu_e}{\mu_o}.$$

But $\sin \rho = AC''/AA' = \sin i/\mu_e$, hence

$$\tan r' = \frac{a}{b} \tan \rho = \frac{a^2 \sin i}{b\sqrt{1 - a^2 \sin^2 i}} = \frac{\mu_o \sin i}{\mu_e \sqrt{\mu_e^2 - \sin^2 i}}.$$

This relation, like the preceding, has been verified by Malus, and it therefore affords additional evidence that the surface of the extraordinary wave is an ellipsoid of revolution.

**180. Negative and Positive Crystals.**—All uniaxal crystals may be divided into two classes. In the first class, to which Iceland spar belongs, the wave surface consists of a spheroid with a sphere interior to it touching it at the extremities of the axis minor. The radius of the sphere is consequently always less than the corresponding radius of the spheroid, and the ordinary velocity is less than the extraordinary. The ordinary index is consequently greater than the extraordinary and the latter ray is less bent towards the normal than the former. For this reason such crystals are called *negative or repulsive crystals.*

In the second class, to which quartz belongs, the reverse is the case, the spheroid is inside the sphere and the ordinary velocity is therefore greater than the extraordinary. The index of refraction of the extraordinary ray is consequently greater than that of the ordinary ray. The former ray is more bent towards the normal than the latter. These are therefore called *positive or attractive crystals.*

Critical angle. Since the sine of the angle of incidence at total reflection is the reciprocal of the refractive index, it follows that the critical angle is always less for the ordinary than for the extraordinary ray in negative crystals, while the reverse is the case in positive crystals.

The following tables contain the values of the ordinary and extraordinary indices for a few crystals :—

POSITIVE CRYSTALS

| | $\mu_o$ | $\mu_e$ |
|---|---|---|
| Quartz | 1·544 | 1·553 |
| Sulphate of Potash | 1·493 | 1·502 |
| Dioptase | 1·667 | 1·723 |
| Ice | 1·306 | 1·307 |
| Zircon | 1·92 to 1·96 | 1·97 to 2·1 |

NEGATIVE CRYSTALS

| | $\mu_o$ | $\mu_e$ |
|---|---|---|
| Iceland Spar | 1·658 | 1·486 |
| Tourmaline | 1·637 to 1·644 | 1·619 to 1·622 |
| Beryl | 1·584 to 1·577 | 1·578 to 1·572 |
| Apatite | 1·646 | 1·642 |
| Nitrate of Soda | 1·5854 | 1·3369 |

**181. Wave Velocity and Ray Velocity.**—In the case of ordinary refraction, such as occurs in isotropic media, the wave diverging from any point is a sphere, and, as shown in Art. 65, the refracted ray is perpendicular to the front of the refracted wave. For this reason the direction of the ray is the same as the direction of propagation of the wave, and the ray velocity is the same as the wave velocity.

The direction of the ray in all cases is found by joining the centre of disturbance A (Fig. 46) to the point C, in which the secondary wave diverging from A touches the wave envelope A′C. When the secondary wave is spherical the ray is perpendicular to the wave envelope, but when the secondary wave has any other shape the ray in general will not be normal to the wave envelope, and the direction of the ray will not be the same as the direction of propagation of the wave. Thus in Fig. 153 A′C is the front of the ordinary wave, and AC is the ordinary ray emanating from A. So also A′C′ is the front of the extraordinary wave—that is, the extraordinary wave envelope, and AC′ is the extraordinary ray from A. The extraordinary ray is therefore in general not at right angles to the front of the extraordinary wave. The ray.

Now as the secondary wave (that is, the sphere and spheroid) diverges from A the point C and C′ move along the lines AC and AC′ with certain definite velocities, termed the *ray velocities*; and it is clear that the ray velocities are proportional to the radii AC and AC′ respectively. At the same time the planes A′C and A′C′ (that is, the wave fronts) move forward with certain definite velocities, termed the *wave velocities*. The direction of wave propagation is normal to the wave front, thus the ordinary wave front moves in the direction of the normal to A′C, while the extraordinary wave moves in a direction perpendicular to A′C′. It is clear, therefore, that the direction of propagation and velocity of the extraordinary wave are in general not the same as that of the extraordinary ray, for while the ray velocity is proportional to the radius vector AC′ the wave velocity is proportional to the perpendicular from A on A′C′. The relations connecting the wave velocities and the ray velocities may be very easily found as follows:—

*Relation between the ordinary and extraordinary wave velocities.*—Since the wave is propagated in the direction of the normal, it is clear that the ordinary wave velocity is measured by AC, the perpendicular from A on A′C, while the velocity of the extraordinary wave is measured by the perpendicular $p$ from A on A′C′. Now since A′C′ is a tangent to the ellipse $x^2/a^2 + y^2/b^2 = 1$, it follows that the perpendicular $p$ from A on A′C′ is given by the equation

$$p^2 = a^2 \sin^2 \alpha + b^2 \cos^2 \alpha,$$

where $\alpha$ is the angle which $p$ makes with the axis minor of the ellipse

—that is, the angle between the wave normal and the optic axis. From this equation we obtain at once the relation

$$p^2 - b^2 = (a^2 - b^2)\sin^2 \alpha.$$

Now $b$ is the velocity of the ordinary wave, and this equation proves that the difference of the squares of the two wave velocities is proportional to the square of the sine of the angle between the optic axis and the direction of wave propagation, or normal to the front of the extraordinary wave.

*Relation between the ordinary and extraordinary ray velocities.*—Since the ray velocities are measured by the radii AC and AC′, then if we denote AC′ by $r$ and the angle which it makes with the optic axis by $\theta$, we have at once from the equation of the ellipse

$$\frac{1}{r^2} = \frac{\sin^2\theta}{a^2} + \frac{\cos^2\theta}{b^2},$$

which gives at once the relation

$$\frac{1}{r^2} - \frac{1}{b^2} = \left(\frac{1}{a^2} - \frac{1}{b^2}\right)\sin^2\theta.$$

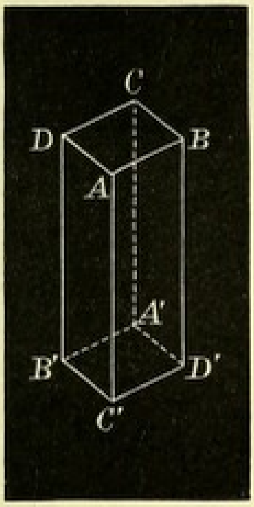
Hence the difference between the squares of the reciprocals of the ordinary and extraordinary ray velocities bears a constant ratio to the square of the sine of the angle between the extraordinary ray and the optic axis.

This proposition is due to Biot, and similar relations of a more general nature will be found to hold for biaxal crystals (see Arts. 205, 206).

**182. Nicol's Prism.**—The most effective and convenient method of procuring a strong beam of plane-polarised light is by double refraction. When ordinary light is transmitted through a crystal of Iceland spar two refracted beams arise, and these we have seen are both plane-polarised and their planes of polarisation are at right angles. Hence if one of the beams is intercepted by any means the other will furnish a pencil of plane-polarised light. An attempt might be made to stop one of the two refracted beams by placing an opaque diaphragm on the second face of the crystal, but it may be easily seen that this method would present difficulties, for unless the source of light is very small or the rhomb very long the refracted beams will overlap.

The former condition entails a great reduction of the illumination, and the latter requires large specimens of Iceland spar of sufficient purity, which are costly. The difficulty might, however, be evaded by receiving the light on a lens placed in contact with the first face of the crystal. After transmission through the lens and crystal the light will converge to two foci—the ordinary rays to one and the extraordinary to the other. One of these foci may now be covered by an opaque diaphragm, and the rays diverging from the other received by a second lens and reduced to parallelism if necessary. The first lens may be plano-convex and may be cemented to the face of the crystal with Canada balsam.

The most convenient method, however, is to stop one of the pencils by total reflection inside the crystal. This is the method adopted in Nicol's prism. A long lozenge-shaped rhomb of calcspar is formed by cleavage from a crystal, so that its length AC′ (Fig. 159) is about three times its width AD. This rhomb is cut in two[1] by a plane passing through the obtuse angles A and A′ and parallel to the longer



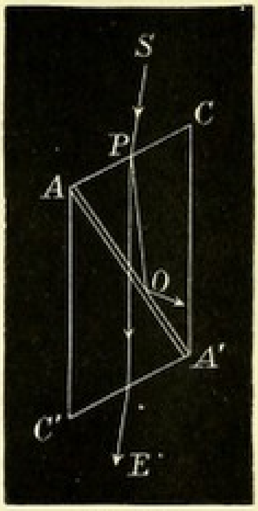




Fig. 159. Nicol's Prism. Fig. 160.

diagonal BD of the end—that is, perpendicular to the principal plane. The cut faces are then polished and cemented together in their original position by a thin film of Canada balsam. The refractive index of the balsam is greater than that of the extraordinary ray in the spar and less than that of the ordinary. Now total reflection occurs only in passing from a more to a less refracting medium. It follows, therefore, that the extraordinary ray falling on the balsam will be always transmitted, but if the incidence be sufficiently oblique the ordinary ray will be totally reflected at the surface of the balsam and refused transmission through the crystal (Fig. 160). The extraordinary ray alone is therefore transmitted, and the light emerging from the prism is plane-polarised at right angles to the principal plane.

The plane of division of the crystal must be drawn so that the

[1] Iceland spar is rather friable, and in practice it is found easier to grind away half of the rhomb instead of cutting it as generally described. The remaining halves of two rhombs thus ground are then cemented together.

light incident nearly normally on the end of the rhomb may fall upon the Canada balsam at an angle not less than the critical angle for the ordinary ray. This angle is easily calculated, for the ordinary index of the spar is about 1·65, while the index of the balsam is nearly 1·55. Hence the index of refraction of the ordinary ray from the crystal to the balsam is $1{\cdot}55/1{\cdot}65 = {\cdot}939$, and the sine of the angle of incidence for total reflection must have this number for its minor limit. Hence if the angle of incidence of the ordinary ray on the balsam is equal to or greater than 69° 30′, total reflection will occur.

**183. Foucault's Prism.**—The Canada balsam in Nicol's prism might be replaced by any substance of less refractive index than calcspar for either ray, or both. It is clear that the less the index of the substance between the two segments of the prism the less the critical angle, and the less the critical angle the shorter the rhomb required to construct a prism of given width.

For this reason Foucault dispensed with the balsam, or cement, altogether, and in the prism which bears his name there is a film of air between the two segments. The critical angles for the ordinary and extraordinary rays are about 37° 14′ and 42° 23′. Hence if the angle of incidence on the film of air is intermediate between the critical angles of the ordinary and extraordinary rays, the former will be totally reflected and the latter transmitted. Although the use of the air film permits of a considerable shortening and consequent reduction in price of the rhomb, yet there is more loss of illumination by reflection from the film. With Nicol's prism the index of the balsam is so near that of the extraordinary ray that the last-named is transmitted almost in its entirety.

**184. Rochon's Prism.**—If a small pencil of light be transmitted through a parallel plate of a doubly refracting substance both the emergent pencils will be parallel to the incident beam and therefore parallel to each other, while their interval of separation will be proportional to the thickness of the plate for a given angle of incidence. But if the faces of the plate be inclined at an angle, so as to form a prism, the emergent beams will be inclined to each other and their separation will increase as they recede from the prism.

Such a separation of the rays is useful in many investigations, and in order to render the divergence as wide as possible the prism should be cut with its refracting edge parallel to the optic axis, so that the incident light may be perpendicular to that axis, for in this case the difference of the ordinary and extraordinary indices is greatest. Such a doubly refracting prism may be achromatised by means of a prism of glass with its refracting edge turned in the opposite direction.

A better arrangement however is that employed by Rochon. Two prisms of calcspar, or quartz, of the same angle are cut so that the refracting edge in one is parallel to the optic axis and in the other perpendicular to it (Fig. 161). They are then cemented together with their edges in opposite directions so as to form a parallelopiped. Thus in a normal cross section of the united prisms the section ABD of one contains the axis, its direction being perpendicular to the face AD, while the section BCD of the other is perpendicular to the axis.

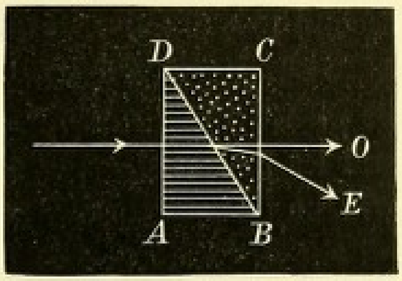


Fig. 161.—Rochon's Prism.

A ray incident normally on the face AD of the first prism travels through it without division, as its direction is parallel to the optic axis; on arriving at the interface double refraction occurs, but the ordinary ray proceeds undeviated. The extraordinary, however, is deviated towards the edge or towards the base of the prism ADB according as the crystal is positive or negative. In the first prism the two rays travel with a common velocity, namely, that of the ordinary ray, and in negative crystals the ordinary velocity is least, while in positive crystals it is the greatest. If A denotes the angle of the prism and $\delta$ the deviation, then since the angle of incidence on the interface is A, the angle of refraction is $A + \delta$, and

$$\frac{\sin(A+\delta)}{\sin A} = \frac{v_e}{v_o} = \frac{a}{b},$$

or, as $\delta$ is small, we have approximately

$$1 + \delta \cot A = a/b,$$

from which we find

$$\delta = \frac{a-b}{b} \tan A.$$

Again, if $r$ be the angle of emergence from the other face of the prism, then, since the angle of incidence there is $\delta$, we have

$$\frac{\sin \delta}{\sin r} = v_e = a,$$

if we take the velocity in air as unity. Consequently $\delta = a \sin r$, and we have, by the previous equation,

$$\sin r = \left(\frac{1}{b} - \frac{1}{a}\right) \tan A = (\mu_o - \mu_e) \tan A,$$

which gives the angular separation ($r$) of the ordinary and extraordinary rays, since the ordinary ray traverses the system normally without deviation.

Rochon's prism is ordinarily constructed of quartz, so that the deviation is in the opposite direction—that is, towards the base of the second prism, since in quartz $\mu_o$ is less than $\mu_e$.

**185. Rochon's Double Image Micrometer.**—If a Rochon's prism be placed in an ordinary telescope (Fig. 162) between the object glass and its principal focus, two images, an ordinary and an extraordinary, will be formed there. The distance between the images will depend on the distance of the prism from the focus; accordingly, by moving the prism suitably, the images may be brought into contact. For this purpose the prism is movable within the telescope, and when the images appear in contact the distance of the prism from the focus can be read off on a graduated scale. By this means Arago determined the apparent diameters of the planets with great accuracy.

Let $f$ be the focal length OF of the object glass and $x$ the distance

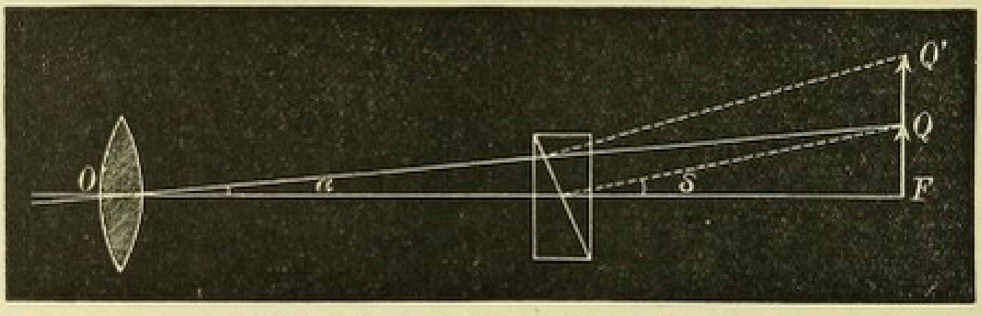



Fig. 162.—Double Image Micrometer.

between the prism and the focus F. Then, if $\alpha$ be the apparent angular diameter of the planet,

$$FQ = f \tan \alpha = x \tan \delta,$$

in which, if we know $f$, $x$, and $\delta$ the deviation, we can determine $\alpha$. For a given instrument the quantities $f$ and $\delta$ will be constant, and could be determined by direct observation on an object of known diameter placed at a known distance. The quantity $\tan \delta / f$ being determined may be regarded as the constant of the instrument.

Again, if an object of height $h$ be at a distance $d$, we have

$$h = d \tan \alpha,$$

from which we find either $h$ or $d$ if the other be known.

**186. Wollaston's Prism.**—This prism differs from that of Rochon only in that the optic axis of the first prism ABD (Fig. 163) is parallel to the face AB, so that it is merely Rochon's prism turned through a right angle. A ray incident normally on the face AB travels along the normal in the crystal as an ordinary ray with velocity $v_o$ and also as an extraordinary ray with velocity $v_e$. On reaching the interface the ordinary ray is refracted extraordinarily, for the principal planes of the two prisms are at right angles, hence the angle of emergence of the ordinary ray is, as in Rochon's prism, given by the equation,

$$\sin r = (\mu_o - \mu_e) \tan A.$$

The extraordinary ray in the first prism ABD traverses it with a velocity $v_e$, but it traverses the second prism as an ordinary ray with a velocity $v_o$. Hence, exactly as before, the angle of emergence from the second face CD is also given by the above equation. The two emerging rays are therefore equally deviated on opposite sides of the normal and their angular separation is doubled. By this duplication the feeble double refraction of quartz is rendered very sensible.

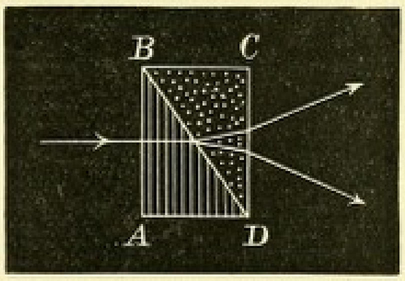


Fig. 163.—Wollaston's Prism.

It should be remarked, however, that the deviation is different for the different colours, and the images are coloured. In Rochon's prism the ordinary image has suffered no deviation and is therefore uncoloured.

# CHAPTER XII

## DOUBLE REFRACTION (FRESNEL'S THEORY)

**187. The Wave Surface in Crystalline Media.** — In homogeneous isotropic substances, such as glass, the physical qualities are the same in all directions around any point, and if any element be displaced the forces of restitution called into play will be opposite and parallel to the displacement. The disturbance will travel with the same velocity in all directions, and the wave surface will be spherical. The optical properties of such substances are also alike in all directions, and the elementary ether waves are spherical.

It is otherwise in the case of crystalline substances. Here the physical properties, such as hardness, elasticity, compressibility, and conductivity for heat and electricity are different in different directions around any point. If any element of such a medium be displaced the forces of restitution will not in general be parallel to the displacement, but may be inclined to it, and will not in general tend to pull the element directly back into its original position. A disturbance will not travel with the same velocity in all directions, and the wave surface will not be spherical.

It is consequently to be expected that the optical character of a crystalline substance will also depend upon its structure, and that the velocity of propagation of the ether waves will be different in different directions, so that the wave diverging from any point will not be a sphere, but a surface of some shape determined by the state of the ether within the crystal and its relation to the molecules of the crystal. In the case of a uniaxal crystal we have already seen that the wave surface consists of two sheets, one a sphere and the other a spheroid.

Thus if we assume that the ether within a crystal possesses the properties of a homogeneous elastic solid, the solution of the problem of double refraction is afforded by the theory of elasticity, and as such it has been developed by Green, MacCullagh, Neumann, and Cauchy. This is known as the elastic solid theory, and is based on the supposition that the ether consists of distinct particles, regarded as material points (at least as far as Fresnel is concerned), which exert forces on each other in the directions of the lines joining them and varying as some function of the distance. In crystalline media the arrangement of the ether particles is supposed to be different in different directions, but symmetrical with respect to three rectangular planes. The ether is thus modified in its arrangement or properties by the presence of the crystalline matter. According to Fresnel the density is modified, while according to Neumann and MacCullagh it is the elasticity which undergoes modification. In all cases the ether is supposed homogeneous, and its axes of symmetry are parallel to those of the crystal.

We might, however (very reasonably), assume the ether to be isotropic everywhere, and the same in all bodies as in free space, and then proceed to explain reflection, refraction, dispersion, etc., as the effects of the momentum communicated to the matter particles by the motion of the ether. When light traverses a transparent substance the matter particles may be set in vibration, thus absorbing some of the energy of ether waves and impeding their progress; and further, the amount of this may depend upon the direction of propagation in crystalline media, but it will be the same in all directions in isotropic substances. Within a crystal the velocity of propagation and the absorption of energy in any direction may also depend upon the relation of that direction to the direction of vibration. It is on this basis of the interaction of the ether and the matter particles that Boussinesq, Voigt, Sellmeyer, Helmholtz, Lommel, and Thomson have built their theories.[1]

Isotropic ether.

**188. The Wave Surface as an Envelope.**—Whatever be the nature of the assumed dynamical conditions, everything is finally reduced to the determination of the velocity of propagation of a plane wave, and the mode of vibration which must exist in such a wave in order that it may be propagated with a determinate velocity. The problem before us is therefore to determine the law according to which a plane wave travels in any assigned direction through a crystal. When this is known the deduction of the form and properties of the wave surface becomes merely a matter of geometry.

Fresnel arrived at the form of the wave surface by considering it as the envelope of a system of plane waves. Thus if a system of plane waves starts from any point in various directions at the same instant, each wave travelling in the direction of its normal with a velocity depending on its direction of propagation, then after any given interval

[1] See Glazebrook's "Report on Optical Theories," *Brit. Assoc. Report*, 1885.

all these plane waves will touch the wave surface, for the surface touching all these planes will be the limit to which a disturbance travelling out in all directions from this centre will have reached in the given time. The construction for the wave surface considered as an envelope is therefore as follows:—On the radii drawn out from a point measure lengths proportional to the velocities of normal propagation in their directions, and at each point thus determined draw a plane perpendicular to the corresponding radius. The envelope of these planes will be the wave surface.

The principle involved in the preceding, viz. that the plane wave moves parallel to itself, is very fundamental. For this wave preserves not only its direction of motion but also the identity of its vibration, and, *for these to be preserved, it is necessary that the elastic forces called into play by the displacement should be parallel to the direction of the displacement.*

**189. Fresnel's Hypotheses.**—The hypotheses on which Fresnel founded his theory may be summarised as follows:—

(1) The vibrations of polarised light are at right angles to the plane of polarisation.

(2) In all cases the elastic forces called into play by the propagation of a train of plane waves (the vibrations being rectilinear and transverse) bear a constant ratio to the elastic forces developed by the displacement of a single molecule, the others remaining at rest.

(3) When a plane wave is propagated in any homogeneous medium, it is only the component of the elastic force parallel to the wave front which is effective in the propagation of the wave.

(4) The velocity of propagation of a plane wave, of permanent type, in any homogeneous medium, is proportional to the square root of the effective component of the elastic force developed by the vibrations.

Little can be said in support of the second hypothesis, and Fresnel himself was conscious of its weakness. Nevertheless the close agreement of experiment with the results of Fresnel's theory must always entitle it to favourable consideration. The fourth hypothesis is introduced on account of a vague analogy between the transverse vibrations of the ether and those of a stretched string, while the difficulty of the third is removed by the supposition that the ether is incompressible, so that the velocity of propagation of the longitudinal vibrations is infinite. Although dynamically unsound Fresnel's theory of double refraction will ever possess a high historic interest, and we shall accordingly detail its leading features and discuss the geometry of the wave surface deduced from it.

**190. The Ellipsoid of Elasticity.**—Fresnel assumed the ether to consist of particles mutually attracting each other, and which, when disturbed from their positions of rest, vibrate under the influence of their mutual attraction.

Assuming the ether to be molecular, and that each molecule is in stable equilibrium under the influence of the others, we can show that the potential energy of displacement of a single molecule is a quadratic function of the three components of the displacement parallel to any set of mutually rectangular axes.

Let V be the potential of the molecule at any point $x$, $y$, $z$, due to the whole system of particles. The components parallel to the axes of reference, of the force on the particle at this point, will be each zero, since it is in equilibrium initially, and therefore we have

$$\frac{dV}{dx}=0, \quad \frac{dV}{dy}=0, \quad \frac{dV}{dz}=0 \qquad (1).$$

Now let the particle at $x$, $y$, $z$ be displaced to a near point $x+\xi$, $y+\eta$, $z+\zeta$, while the others remain at rest. The potential at this point will be $V+dV$, or, since $\xi$, $\eta$, $\zeta$ are supposed very small,

$$V+\xi\frac{dV}{dx}+\eta\frac{dV}{dy}+\zeta\frac{dV}{dz}+\tfrac{1}{2}\left(\xi^2\frac{d^2V}{dx^2}+\eta^2\frac{d^2V}{dy^2}+\zeta^2\frac{d^2V}{dz^2}+2\eta\zeta\frac{d^2V}{dydz}+\text{etc.}\right) \quad (2).$$

The components of the force on the molecule are found by differentiating (2) with respect to $\xi$, $\eta$, $\zeta$ respectively. Hence, remembering the equations (1), we have

$$\left.\begin{aligned} X&=\xi\frac{d^2V}{dx^2}+\eta\frac{d^2V}{dxdy}+\zeta\frac{d^2V}{dxdz}\\ Y&=\xi\frac{d^2V}{dydx}+\eta\frac{d^2V}{dy^2}+\zeta\frac{d^2V}{dydz}\\ Z&=\xi\frac{d^2V}{dzdx}+\eta\frac{d^2V}{dzdy}+\zeta\frac{d^2V}{dz^2}\end{aligned}\right\} \qquad (3),$$

or, writing

$$\frac{d^2V}{dx^2}=A, \quad \frac{d^2V}{dy^2}=B, \quad \frac{d^2V}{dz^2}=C, \quad \frac{d^2V}{dydz}=F, \quad \frac{d^2V}{dzdx}=G, \quad \frac{d^2V}{dxdy}=H,$$

we have, for the forces parallel to the axes of reference,

$$\left.\begin{aligned} X&=A\xi+H\eta+G\zeta\\ Y&=H\xi+B\eta+F\zeta\\ Z&=G\xi+F\eta+C\zeta\end{aligned}\right\} \qquad (4).$$

Hence if we construct the quadric

$$A\xi^2+B\eta^2+C\zeta^2+2F\eta\zeta+2G\zeta\xi+2H\xi\eta=1 \qquad (5),$$

we have, denoting the left-hand member of this equation by 2S,

$$X=\frac{dS}{d\xi}, \quad Y=\frac{dS}{d\eta}, \quad Z=\frac{dS}{d\zeta} \qquad (6).$$

Now $\xi$, $\eta$, $\zeta$ are proportional to the direction cosines $\alpha$, $\beta$, $\gamma$, of the displacement, and if the displacement be taken as unity, $\xi$, $\eta$, $\zeta$ will be numerically equal to $\alpha$, $\beta$, $\gamma$ respectively; but the resultant force on the molecule will not be in general in the same direction as the displacement, for the direction cosines of the resultant force are proportional to X, Y, Z, and these by (6) are proportional to the direction cosines of the normal to the quadric (5) at the point $\xi$, $\eta$, $\zeta$. The resultant force on the molecule is consequently not parallel to the direction of displacement, viz. along the radius vector to $\xi$, $\eta$, $\zeta$, but is along the normal to the quadric.

There are, however, three directions at any point along which if a molecule be displaced the resultant force will be parallel to the displacement, and tend to restore the particle to its original position. These directions are the axes of the quadric (5), and if we take them for axes of co-ordinates the equation of the quadric may be written in the form

$$a^2x^2 + b^2y^2 + c^2z^2 = 1 \qquad (7).$$

In this case the expression (4) for the components of the force on a molecule displaced through a distance $\rho = \sqrt{\xi^2 + \eta^2 + \zeta^2}$ will become

$$\left.\begin{array}{l} X = a^2\xi, \quad Y = b^2\eta, \quad Z = c^2\zeta \\ X = \rho a^2\alpha, \quad Y = \rho b^2\beta, \quad Z = \rho c^2\gamma \end{array}\right\} \qquad (8),$$

or (second line above)

while the restoring force along the line $\alpha$, $\beta$, $\gamma$ *of the displacement* will be

$$F = X\alpha + Y\beta + Z\gamma = \rho(a^2\alpha^2 + b^2\beta^2 + c^2\gamma^2),$$

or

$$F = \frac{\rho}{r^2} \qquad (9),$$

where $r$ is the radius vector of the quadric (7) drawn in the direction $\alpha$, $\beta$, $\gamma$.

Hence if we consider only the component F as effective, the equation of motion of the particle will be

$$\frac{d^2\rho}{dt^2} = -\frac{\rho}{r^2} \qquad (10),$$

and the time of vibration will consequently be given by the equation

$$T = 2\pi r \qquad (11).$$

But the velocity of propagation is connected with the wave length, and periodic time, by the equation $\lambda = vT$, therefore

$$v = \frac{\lambda}{2\pi r} \qquad (12),$$

or the velocity of propagation of a plane wave, in which the displacement of the molecules is parallel to a given direction, is inversely proportional to the radius vector $r$ of the ellipsoid of elasticity drawn in that direction.

In arriving at this result it will be seen that the assumptions used are, that the force on any particle depends on the *absolute* displacement from the original position, and is the same as if the other particles remained at rest, whereas in any kind of wave motion the true force on a particle should depend on its displacement *relatively* to the others. So also it is only the *component* force along the line of displacement that has been considered effective in the equation of motion (10) of the particle. Assumptions.

*Cor.*—If $v$ be the velocity of propagation of a plane wave whose vibrations are in the direction $\alpha$, $\beta$, $\gamma$, then by (12), since $v \propto 1/r$, we may write equation (7) in the form,

$$v^2 = v_1^2\alpha^2 + v_2^2\beta^2 + v_3^2\gamma^2 \qquad (13),$$

or

$$(v^2 - v_1^2)\alpha^2 + (v^2 - v_2^2)\beta^2 + (v^2 - v_3^2)\gamma^2 = 0 \qquad (14),$$

where $v_1$, $v_2$, $v_3$ are now the velocities of propagation of waves vibrating parallel to the axes of elasticity. If $\mu_1$, $\mu_2$, $\mu_3$ be the principal refractive indices we may write, if we take the velocity outside the crystal to be unity,

$$\mu_1 = \frac{1}{v_1}, \quad \mu_2 = \frac{1}{v_2}, \quad \mu_3 = \frac{1}{v_3} \qquad (15),$$

and (13) takes the form

$$v^2 = \frac{\alpha^2}{\mu_1^2} + \frac{\beta^2}{\mu_2^2} + \frac{\gamma^2}{\mu_3^2} \qquad (16).$$

**191. Singular Directions.**—The essential condition, for the propagation of a plane wave without alteration, is that the effective component of the elastic force developed by the displacement should be parallel to the displacement. This condition is only satisfied for two directions in any plane, viz. the axes of the conic in which that plane cuts the preceding quadric.

When the elements in the front of a plane wave are displaced parallel to a given direction in the wave front, we have seen that the force of restitution on each element will not in general be in the direction of the displacement, nor even in the plane, but will be normal to the ellipsoid of elasticity—that is, perpendicular to the central section of the ellipsoid which is conjugate to the direction of displacement. Thus if AB (Fig. 164) be the section of the ellipsoid of elasticity by the plane of the wave front, and OA the direction of displacement of the elements in the front of the wave, OB the radius of the section conjugate to OA, and OC the radius of the ellipsoid conjugate to the section AB, then the force of restitution will be parallel to ON the normal to the plane BOC. Now if the projection of ON on the plane AOB coincides with OA, the plane of ON and OA must be perpendicular to the plane AOB. But ON is perpendicular to OB, therefore OB will be perpendicular to OA—that is, OA and OB are the axes of the section AOB. Thus the resultant force on any particle may be resolved into two components, one in the wave front and the other perpendicular to it. If the displacement is along either of the axes of the elliptic section AOB, the force component in the wave front will also be along that axis, and these axes are the only directions in the wave front possessing this property. These two directions are termed the *singular directions* in the plane of the wave.

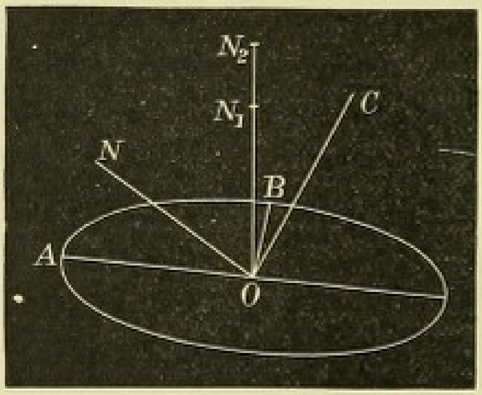


Fig. 164.

Circular sections.

If the plane be parallel to either of the circular sections of the quadric, every direction is a singular direction, and a vibration in any direction in this plane will be propagated without alteration and there will consequently be no double refraction.

Thus when a displacement occurs in any direction in the front of a wave it is only its components parallel to the singular directions that are propagated as permanent waves, and these are propagated with different velocities (except when the plane is parallel to a circular section). Consequently the bifurcation of the ray on entering a crystal is accounted for, as is also the polarisation of the two rays, and the fact that their planes of polarisation are at right angles.

Polarisation.

We have now arrived at the fundamental law of double refraction—

*In one and the same direction two systems of plane waves are propagated normally, having their vibrations parallel to the axes of the section of the ellipsoid of elasticity by a diametral plane perpendicular to the direction, and the velocities of normal propagations of the two systems are inversely proportional to the lengths of these axes.*

We see also that there are in general two directions of propagation along which there will be no double refraction, and these directions are perpendicular to the circular sections of the ellipsoid of elasticity. They are termed the axes of single wave velocity or the *optic axes* of the crystal. For waves propagated in these directions the velocity is the same whatever be the direction of vibration. The wave fronts will coincide, but there may be a separation of the rays (Art. 202).

ptic axes.

**192. Problem.**—To a variable plane drawn through the centre of the ellipsoid

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}+\frac{z^2}{c^2}=1$$

a perpendicular is drawn at the centre of the section, and portions $ON_1$ and $ON_2$ (Fig. 164) are taken on it which, measured from the centre, are equal to the axes of the section, to find the locus of their extremities.

If $r$ be the length of either axis, and $l$, $m$, $n$ the direction cosines of the normal to the plane, we have [1]

$$\frac{a^2l^2}{r^2-a^2}+\frac{b^2m^2}{r^2-b^2}+\frac{c^2n^2}{r^2-c^2}=0.$$

(See Salmon's *Geometry of Three Dimensions*, Art. 101.)

In this equation $r$ is the length of the radius vector of the required locus, and $l$, $m$, $n$ are its direction cosines; the equation of the surface is therefore

$$\frac{a^2x^2}{r^2-a^2}+\frac{b^2y^2}{r^2-b^2}+\frac{c^2z^2}{r^2-c^2}=0.$$

**193. The Normal Velocity Surface.**—Around any point O construct the ellipsoid of elasticity, and consider a system of plane waves passing through O in all directions at the same instant. Let any one of the planes cut the ellipsoid in the section AOB (Fig. 164) of which the axes are OA and OB. Draw a normal at O to the plane of the section and measure off $ON_1$ and $ON_2$ on it inversely proportional to the axes OA and OB, then if planes be drawn through $N_1$ and $N_2$ parallel to the plane of the section, they will represent simultaneous

[1] If $a'$, $b'$ be the axes of any central section, $l$, $m$, $n$ the direction cosines of the normal to its plane, and R the intercept made on it by the surface, we have

$$\frac{1}{a'^2}+\frac{1}{b'^2}+\frac{1}{R^2}=\frac{1}{a^2}+\frac{1}{b^2}+\frac{1}{c^2},$$

therefore

$$\frac{1}{a'^2}+\frac{1}{b'^2}=\frac{1}{a^2}+\frac{1}{b^2}+\frac{1}{c^2}-\frac{l^2}{a^2}-\frac{m^2}{b^2}-\frac{n^2}{c^2}.$$

But

$$\frac{1}{a'^2b'^2}=\frac{p^2}{a^2b^2c^2}=\frac{l^2}{b^2c^2}+\frac{m^2}{c^2a^2}+\frac{n^2}{a^2b^2}.$$

Whence the quadratic for $r$, either semi-axis, is

$$\frac{1}{r^4}-\frac{1}{r^2}\left(\frac{1-l^2}{a^2}+\frac{1-m^2}{b^2}+\frac{1-n^2}{c^2}\right)+\frac{l^2}{b^2c^2}+\frac{m^2}{c^2a^2}+\frac{n^2}{a^2b^2}=0,$$

which may be written in the form above, by remembering that $1=l^2+m^2+n^2$.

positions of the waves which will be propagated with their fronts normal to $ON_1N_2$, the one having its vibrations parallel to OA and the other parallel to OB, for we have already seen that the velocities of propagation of these waves are inversely proportional to OA and OB respectively. These planes therefore envelop the wave surface.

If the plane AOB be supposed to turn round O, the points $N_1$ and $N_2$ determined as above will each trace out a sheet of a surface, and the locus of both will therefore be a surface of two sheets, termed the surface of *normal velocities*, since any radius vector of it determines the two normal velocities of the plane waves propagated in that direction.

The equation of this surface is easily found, for it is described as the locus of points on the normal (at the centre) to a variable central section of the quadric

$$a^2x^2+b^2y^2+c^2z^2=1,$$

the distances from the centre to the points being inversely proportional to the axes of the section. The locus is therefore found from the equation of the preceding article, by changing $a$, $b$, $c$, $r$ into their reciprocals, and we thus obtain

$$\frac{x^2}{r^2-a^2}+\frac{y^2}{r^2-b^2}+\frac{z^2}{r^2-c^2}=0.$$

*Cor.* 1.—If $v$ be the velocity of propagation in the direction $l$, $m$, $n$, we have $r$ proportional to $v$ and $x$, $y$, $z$, proportional to $l$, $m$, $n$, therefore

$$\frac{l^2}{v^2-v_1^2}+\frac{m^2}{v^2-v_2^2}+\frac{n^2}{v^2-v_3^2}=0.$$

*Cor.* 2.—The direction cosines of the vibration being $\alpha$, $\beta$, $\gamma$, we have (Art. 190, Cor.)

$$(v^2-v_1^2)\alpha^2+(v^2-v_2^2)\beta^2+(v^2-v_3^2)\gamma^2=0,$$

and, since the vibration is perpendicular to the wave normal, we have also

$$l\alpha+m\beta+n\gamma=0.$$

Combining these equations with that of Cor. 1 we find

$$\frac{v^2-v_1^2}{l}\alpha=\frac{v^2-v_2^2}{m}\beta=\frac{v^2-v_3^2}{n}\gamma.$$

**194. Equation of the Wave Surface.**—If $v$ be the velocity of wave propagation in the direction $l$, $m$, $n$, the wave surface is the envelope of the plane

$$lX+mY+nZ=v \qquad (1),$$

where $v$ is related to $l$, $m$, $n$ by the equation

$$\frac{l^2}{v^2 - v_1^2} + \frac{m^2}{v^2 - v_2^2} + \frac{n^2}{v^2 - v_3^2} = 0 \qquad (2),$$

and in addition we have always

$$l^2 + m^2 + n^2 = 1 \qquad (3).$$

Differentiating (1), (2), (3), regarding $l$, $m$, $n$ as variables, we obtain [1]

$$Xdl + Ydm + Zdn - dv = 0,$$

$$\frac{ldl}{v^2 - v_1^2} + \frac{mdm}{v^2 - v_2^2} + \frac{ndn}{v^2 - v_3^2} - \left\{ \frac{l^2}{(v^2 - v_1^2)^2} + \frac{m^2}{(v^2 - v_2^2)^2} + \frac{n^2}{(v^2 - v_3^2)^2} \right\} vdv = 0,$$

$$ldl + mdm + ndn = 0.$$

From which, by the use of indeterminate multipliers A and B, we have

$$X = Al + Bl/(v^2 - v_1^2) \qquad (4),$$

$$Y = Am + Bm/(v^2 - v_2^2) \qquad (5),$$

$$Z = An + Bn/(v^2 - v_3^2) \qquad (6),$$

$$Bv \left\{ \frac{l^2}{(v^2 - v_1^2)^2} + \frac{m^2}{(v^2 - v_2^2)^2} + \frac{n^2}{(v^2 - v_3^2)^2} \right\} = 1 \qquad (7).$$

Multiplying (4), (5), (6) by $l$, $m$, $n$ respectively and adding we find

$$v = A \qquad (8),$$

while the same equations, squared and added, give, with the aid of (7),

$$X^2 + Y^2 + Z^2 = A^2 + B/v,$$

or, writing $R^2 = X^2 + Y^2 + Z^2$ and using (8) we obtain

$$B = v(R^2 - v^2) \qquad (9).$$

Substituting these values of A and B in (4) we have

$$X = lv + lv\frac{R^2 - v^2}{v^2 - v_1^2} = lv\frac{R^2 - v_1^2}{v^2 - v_1^2}.$$

Therefore

$$l = \frac{v^2 - v_1^2}{R^2 - v_1^2} \cdot \frac{X}{v}.$$

Similarly

$$m = \frac{v^2 - v_2^2}{R^2 - v_2^2} \cdot \frac{Y}{v},$$

and

$$n = \frac{v^2 - v_3^2}{R^2 - v_3^2} \cdot \frac{Z}{v}.$$

---

[1] Archibald Smith, *Phil. Mag.* 1838, p. 335.

Hence by substituting these values of $l$, $m$, $n$ in (1), we have finally

$$X^2\frac{v^2-v_1^2}{R^2-v_1^2}+Y^2\frac{v^2-v_2^2}{R^2-v_2^2}+Z^2\frac{v^2-v_3^2}{R^2-v_3^2}=v^2=v^2\left(\frac{X^2}{R^2}+\frac{Y^2}{R^2}+\frac{Z^2}{R^2}\right),$$

or

$$\frac{v_1^2X^2}{R^2-v_1^2}+\frac{v_2^2Y^2}{R^2-v_2^2}+\frac{v_3^2Z^2}{R^2-v_3^2}=0,$$

which is the equation of the wave surface.

**195. Direction of Vibration at any Point of the Wave Front.**—We have seen that in each direction in the crystal two waves can in general be propagated with different velocities, and these are plane-polarised in planes at right angles. The directions of the vibrations in these waves are parallel to the axes of the elliptic section of the elasticity ellipsoid. We shall now show that this direction is parallel to the projection of the radius vector of the wave surface on the tangent plane at the corresponding point. For by Cor. 2, Art. 193, we have $l$, $m$, $n$ related to $\alpha$, $\beta$, $\gamma$ by equations of the form

$$\frac{l}{v^2-v_1^2}=\kappa\alpha,\quad \frac{m}{v^2-v_2^2}=\kappa\beta,\quad \frac{n}{v^2-v_3^2}=\kappa\gamma.$$

Therefore the equations (4), (5), (6) of the preceding Article become

$$X=Al+B\kappa\alpha,\quad Y=Am+B\kappa\beta,\quad Z=An+B\kappa\gamma,$$

which show that the direction of vibration $\alpha$, $\beta$, $\gamma$ lies in the plane containing the radius vector to $x$, $y$, $z$, and the perpendicular $l$, $m$, $n$ on the corresponding tangent plane to the wave surface. But the vibration takes place in the wave front, and is therefore parallel to the line joining $x$, $y$, $z$ to the foot of the perpendicular on the tangent plane, or the direction of vibration in any ray is parallel to the projection of the ray on the corresponding tangent plane.

**196. Relation of the Planes of Polarisation to the Optic Axes.**—The planes of polarisation of the two rays which correspond to any given plane wave front are very simply related to the optic axes. Let ON (Fig. 165) be the normal to the plane of the wave, OM and OM′ the optic axes. Then the planes of polarisation of the two rays are the planes which pass through ON, and the axes (OA and OB) of the section in which the ellipsoid of elasticity is cut by the plane through O to which ON is normal. Now the circular sections of the ellipsoid are perpendicular to OM and OM′,

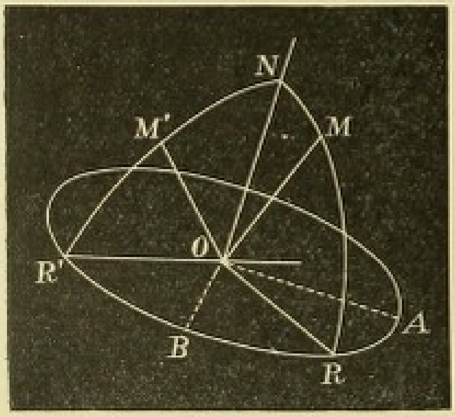


Fig. 165.

they will therefore meet the elliptic section ARB in radii which are equal to each other and perpendicular to the planes MON and M′ON respectively. That is, the radii of the section which are perpendicular to the radii OR and OR′ are equal, OR and OR′ are therefore also equal, and they are consequently equally inclined to the axes OA and OB of the section, and hence the planes containing ON and these axes will bisect the angles between the planes MON and M′ON. That is, the planes of polarisation of the two rays bisect the angles between the planes containing the wave normal and the optic axes.

**197. Equation of the Wave Surface by Means of the Reciprocal Ellipsoid.**—We have already deduced the wave surface as the envelope of a variable plane. It occurred to MacCullagh[1] that the wave surface might also be described as the locus of a point by using the reciprocal quadric

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}+\frac{z^2}{c^2}=1.$$

If any plane be drawn through the centre of this quadric cutting it in a section of axes $a'$ and $b'$, and if along the normal at its centre portions be measured off equal to the axes $a'$ and $b'$ respectively, the locus of the extremities of these portions will be the wave surface. Its equation is thus found at once to be

$$\frac{a^2x^2}{r^2-a^2}+\frac{b^2y^2}{r^2-b^2}+\frac{c^2z^2}{r^2-c^2}=0.$$

Developed in terms of the co-ordinates, $x$, $y$, $z$, the equation becomes

$$(x^2+y^2+z^2)(a^2x^2+b^2y^2+c^2z^2)-a^2(b^2+c^2)x^2-b^2(c^2+a^2)y^2-c^2(a^2+b^2)z^2+a^2b^2c^2=0.$$

The surface is consequently of the fourth degree, and consists of two sheets. We shall suppose $a>b>c$.

**198. Uniaxal Crystal.**—If two of the principal velocities are equal, $b=c$ suppose, the above equation may be written in this form

$$(x^2+y^2+z^2-b^2)[a^2x^2+b^2(y^2+z^2)-a^2b^2]=0.$$

The surface consequently breaks up into the sphere

$$x^2+y^2+z^2=b^2,$$

and the spheroid

$$a^2x^2+b^2(y^2+z^2)=a^2b^2.$$

In this case the sphere and ellipsoid of revolution touch each other where the axis $b$ meets them, and this is the optic axis of the crystal.

[1] MacCullagh, "On the double refraction of light in a crystallised medium according to the principles of Fresnel" (*Trans. Royal Irish Academy*, June 1830).

The sphere lies entirely within the spheroid, so that the ordinary index of refraction is greater than the extraordinary, and the crystal is negative like Iceland spar when $b$ is less than $a$.

If, on the other hand, $b$ is equal to $a$, then the radius of the sphere is equal to $a$, and the spheroid lies entirely within the sphere, the two touching where $a$ meets them; $a$ is the optic axis and the crystal is positive or attractive, the ordinary index being less than the extraordinary.

Finally if $a = b = c$ the wave surface reduces to the sphere

$$(x^2 + y^2 + z^2 - a^2)^2 = 0,$$

which represents the case of isotropic media, or crystals belonging to the cubic system.

**199. Principal Sections of the Wave Surface.**—If in the equation of the wave surface we make successively $x = 0$, $y = 0$,

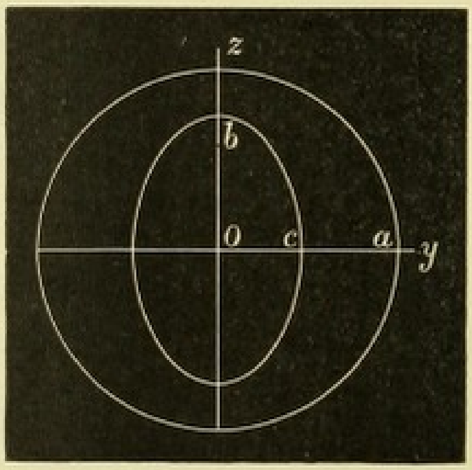



Fig. 166.

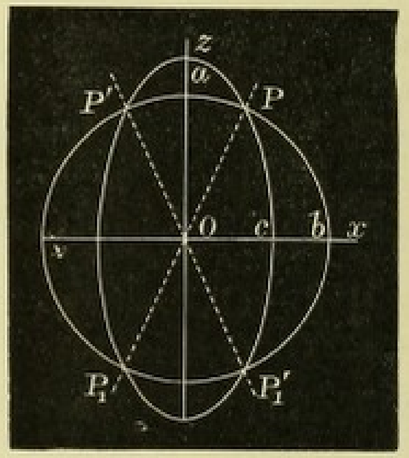



Fig. 167.

$z = 0$ we obtain the equations of the curves of section of the wave surface by the planes of $yz$, $zx$, and $xy$ respectively. Each of these sections consists of two curves, namely, a circle and an ellipse having the same centre.

(1) Thus making $x = 0$, we find

$$(y^2 + z^2 - a^2)(b^2y^2 + c^2z^2 - b^2c^2) = 0,$$

therefore the section of the surface by the plane $yz$ consists of the circle of radius $a$,

$$y^2 + z^2 = a^2,$$

and the ellipse

$$b^2y^2 + c^2z^2 = b^2c^2$$

of axes $b$ and $c$, which consequently lies entirely within the circle, as shown in Fig. 166.

(2) Now suppose $y=0$, and we find the section by the plane $zx$ to consist of the circle

$$x^2+z^2=b^2,$$

and the ellipse

$$a^2x^2+c^2z^2=a^2c^2,$$

the radius of the circle being $b$ and the axes of the ellipse being $a$ and $c$ it follows that the circle meets the ellipse in four points, as represented in Fig. 167.

(3) Again making $z=0$, we find the circle

$$x^2+y^2=c^2,$$

and the ellipse

$$a^2x^2+b^2y^2=a^2b^2.$$

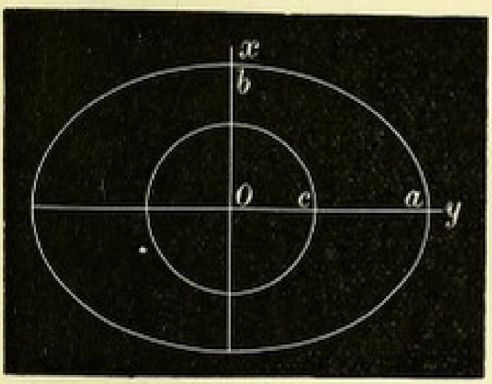



Fig. 168.

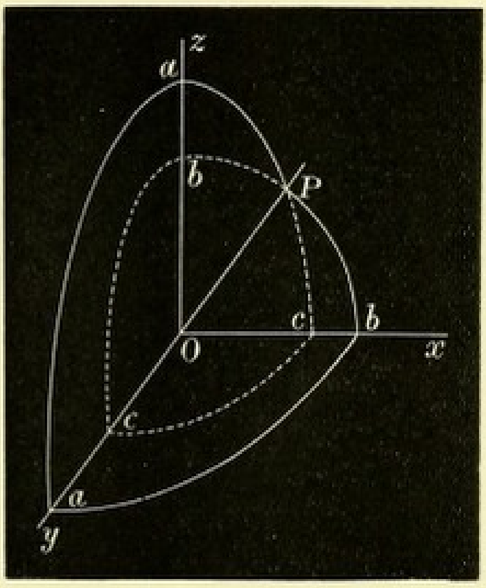



Fig. 169.

The circle therefore lies entirely within the ellipse, as shown in Fig. 168.

Fig. 169 represents a segment of the surface of the wave by the principal planes. It intercepts segments $b$ and $c$ on the axis of $x$, $c$ and $a$ on the axis of $y$, and $a$ and $b$ on the axis of $z$. It presents four conical points in the plane $zx$, where the circle of radius $b$ meets the ellipse of axes $a$ and $c$. The lines OP, OP′ to these points are such that only one ray will be propagated in their direction, and they are consequently called the *axes of single ray velocity*.

Models of the wave surface may be procured, and an examination of one of them will assist the ideas of the student with reference to the nature of the surface.

**200. Construction of Huygens.**—The form of the wave surface being known, the directions of the refracted rays are determined by tangent planes drawn to the two sheets of the surface according to the construction of Huygens. Thus when a plane wave is incident on the face of a biaxal crystal each point of the surface is a centre from which elementary wave surfaces diverge; each wavelet consists of two sheets, and the envelope consists of two planes, one touching all the interior sheets, the other all the exterior. To determine the directions of these planes it is only necessary to construct one wave surface, and draw tangent planes to both its sheets through the trace of the incident wave on the face of the crystal, as has been already indicated in Arts. 66 and 178. The line joining the centre of any wavelet to its point of contact with the tangent plane gives the direction of the refracted ray.

It may happen that no real tangent plane can be drawn to one or both of the sheets of the wave surface under the prescribed conditions, and total reflection will then occur as in the case of isotropic media.

**201. The Optic Axes or Axes of Single Wave Velocity—Axes of Internal Conical Refraction.**—The form of the section of the wave surface by each of the principal planes has been arrived at in Art. 199. Each principal plane cuts the surface in a circle and an ellipse having the same centre, but in only one case, that of the plane $zx$, do the circle and ellipse intersect. Here the radius of the circle is $b$ and the axes of the ellipse are $a$ and $c$. The curves consequently intersect in four points P, P′, $P_1$, $P'_1$, and have four common tangents MN, $M_1N_1$, M′N′, $M'_1N'_1$ (Fig. 170).

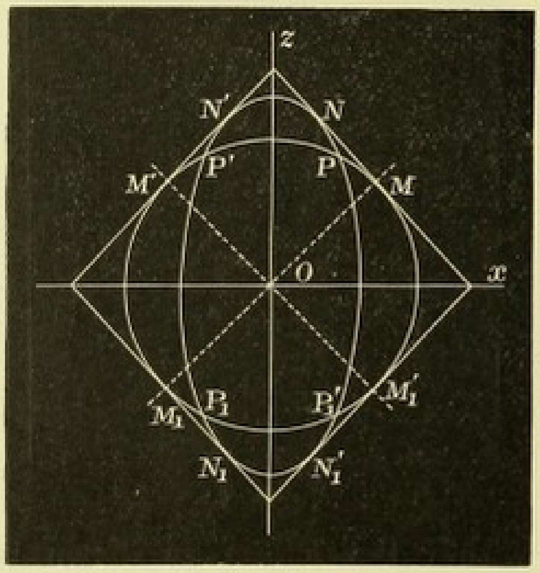


Fig. 170.

Now the planes passing through these common tangents and perpendicular to the plane of the section are tangent planes to the wave surface. Moreover they do not, like ordinary tangent planes to a surface, merely touch it at one point, or even at two points, M and N, etc. The points P, $P_1$, etc., are what are termed conical points on the surface; they are little pits or dimples, and the tangent planes MN, etc., cover them over and touch the surface, as Sir William Hamilton proved, all round the perimeter of a circle of contact.[1]

[1] The points on the surface S at which the tangent plane is parallel to the axis of $y$ satisfy the condition $\frac{dS}{dy}=0$. Applying this to the wave surface we find

$$y\{b^2(x^2+y^2+z^2)-b^2(c^2+a^2)+a^2x^2+b^2y^2+c^2z^2\}=0.$$

The factor $y=0$ corresponds to points situated in the plane $xz$, which obviously

The lines OM, OM′ are perpendicular to these planes, and they are therefore such directions of the wave normal that only one wave front exists, for the plane MN touches both sheets of the wave surface.

The directions OM and OM′ are for this reason termed the axes of *single wave velocity*. For these directions there is only one wave envelope, for other directions there are two; they are therefore the *optic axes* of the crystal.

The angle between the optic axes may be easily expressed in terms of the principal velocities, for since OM $(=b)$ is a perpendicular to the tangent MN to the ellipse $x^2/c^2 + z^2/a^2 = 1$ (Art. 199), its length is given by the equation

$$b^2 = c^2 \cos^2\phi + a^2 \sin^2\phi,$$

where $\phi$ is the angle OM makes with the axis of $x$—that is, half the angle $\text{MOM}'_1$.

Consequently

$$\sin\phi = \sqrt{\frac{b^2 - c^2}{a^2 - c^2}}, \quad \cos\phi = \sqrt{\frac{a^2 - b^2}{a^2 - c^2}}, \quad \tan\phi = \sqrt{\frac{b^2 - c^2}{a^2 - b^2}}$$

or, in terms of the principal indices of refraction,

$$\sin\phi = \frac{\mu_1}{\mu_2}\sqrt{\frac{\mu_2^2 - \mu_3^2}{\mu_1^2 - \mu_3^2}}, \quad \cos\phi = \frac{\mu_3}{\mu_2}\sqrt{\frac{\mu_1^2 - \mu_2^2}{\mu_1^2 - \mu_3^2}}, \quad \tan\phi = \frac{\mu_1}{\mu_3}\sqrt{\frac{\mu_2^2 - \mu_3^2}{\mu_1^2 - \mu_2^2}}.$$

Since $\tan\phi = z/x$ it follows that the equations of the optic axes OM and OM′ are

$$z = \pm\sqrt{\frac{b^2 - c^2}{a^2 - b^2}}\cdot x, \quad \text{and } y = 0,$$

which shows that they are normal to the circular sections of the quadric

$$a^2x^2 + b^2y^2 + c^2z^2 = 1.$$

---

satisfy the conditions of the problem. The second factor represents an ellipsoid which clearly possesses the same planes of circular section as the ellipsoid of elasticity $a^2x^2 + b^2y^2 + c^2z^2 = 1$. Hence a plane wave parallel to a circular section of the ellipsoid of elasticity—that is, perpendicular to either optic axis—will meet the above quadric in a circle, at every point of which it touches the wave surface. For on eliminating $y^2$ between this quadric equation and that of the wave surface, the result breaks up into factors

$$\left(z + x\sqrt{\frac{a^2 - b^2}{b^2 - c^2}} + b\sqrt{\frac{a^2 - c^2}{b^2 - c^2}}\right)\left(z - x\sqrt{\frac{a^2 - b^2}{b^2 - c^2}} + b\sqrt{\frac{a^2 - c^2}{b^2 - c^2}}\right)$$
$$\left(z + x\sqrt{\frac{a^2 - b^2}{b^2 - c^2}} - b\sqrt{\frac{a^2 - c^2}{b^2 - c^2}}\right)\left(z - x\sqrt{\frac{a^2 - b^2}{b^2 - c^2}} - b\sqrt{\frac{a^2 - c^2}{b^2 - c^2}}\right)$$

which are the four tangent planes at the conical points.

In general $a, b, c$ are functions of the wave length, and consequently the angle between the optic axes varies with the colour of the light.

**202. Internal Conical Refraction—Lloyd's Experiment.**—The direction of a refracted ray is given by the line joining the centre of disturbance to the point of contact of the wave surface with the wave envelope, as determined by Huygens's construction.

Hence if O (Fig. 170) be the centre of disturbance (say a point on the face of a crystal on which a plane wave is incident), and if MN be the direction of the front of the refracted plane wave, it follows that any line from O to any point of the circle of contact of MN with the wave surface is a possible direction for the ray in the crystal. We should expect then that a ray incident on the face of the crystal in such a direction that the refracted wave in the crystal is parallel to MN (that is, the plane wave in the crystal travels in the direction of the optic axis) should on entering the crystal be divided not into two rays but into a cone of rays, viz. the cone joining O to the circle of contact of MN with the wave surface.

This result was predicted by Sir William Hamilton, and at his request the experiment was undertaken by Dr. Lloyd,[1] who found the anticipations of the theory verified in a most remarkable manner.

A plate of aragonite was used, having its faces perpendicular to the bisectors of the angle between the optic axes, which, in the crystal submitted to experiment, was about 20°. One of these lines is parallel to OM (Fig. 171), and its direction was determined beforehand by means of the phenomena of the colours of crystalline plates. A slender pencil of light, SO, limited by two screens, one of which, CD, was at some distance from the plate, and the other, which was a thin leaf of metal, was pierced by a small hole and placed on the face of the plate. The emergent rays were received on a screen of silver paper, EF. The minuteness of this phenomenon and the perfect accuracy required in the incidence rendered it very difficult to observe. The crystal was moved with extreme slowness so as to vary the direction of incidence very gradually, and when the required position was obtained the two images on the screen EF suddenly spread out into a continuous ring of light. No sensible enlargement of this ring could be observed as the distance of the screen EF from the plate was altered,

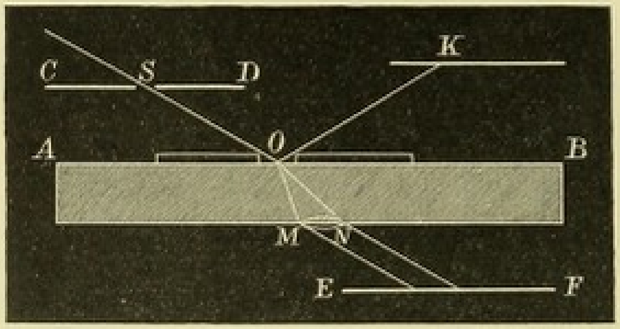


Fig. 171.

[1] Lloyd, *Trans. Roy. Irish Acad.* vol. xvii. p. 145, 1833.

showing that the emergent beam was cylindrical, and that consequently the path of the light in the crystal was the cone OMN. The angle of this cone was found to be $1° 50'$, and its magnitude as indicated by the theory was $1° 55'$, so that the observed and theoretical values agreed very closely.

To measure the angle of incidence Lloyd received the pencil reflected at O on a screen, and marked the point K where it fell. He then removed the plate and arranged a theodolite so that its axis of rotation passed through O. He was then able to measure the angle SOK, which is double the angle of incidence. The observed angle of incidence was $15° 40'$, and its value as indicated by theory $15° 19'$. The agreement of the theory and experiment is thus exceedingly complete. The diameter of the ring on the screen EF determines the angle of the refracted cone.

The existence of conical refraction has been regarded as one of the most striking proofs of the general correctness of Fresnel's theory of double refraction, but Stokes[1] has pointed out that it is not competent to decide between the several theories which lead to Fresnel's wave surface as a near approximation. Internal conical refraction depends upon the existence of a tangent plane to the wave surface which touches it along a plane curve, and this property would be possessed by the wave surface arising from any reasonable hypothesis. Other forms of the wave theory, based on very different assumptions, lead to Fresnel's wave surface exactly. The existence of conical refraction cannot therefore be regarded as deciding in favour of Fresnel's particular doctrine.

*Direction of Vibration.*—It is easy to determine the direction of the vibrations in each of the rays which constitute the cone OMN, for the direction of vibration of any refracted ray is found by projecting the ray on the corresponding tangent plane to the wave surface (Art. 195). Now M is a point on the circle in which the tangent plane touches the wave, and OM is perpendicular to this plane, therefore if ON (Fig. 172) is any ray of the cone its projection on the plane of the circle will pass through M and consequently be the chord MN of the circle. Hence the directions of the vibrations of the different rays of the cone are parallel to the chords of the circle of contact drawn from M to the points where the rays meet this circle. It follows therefore that two rays meeting this circle at diametrically opposite points are such that their

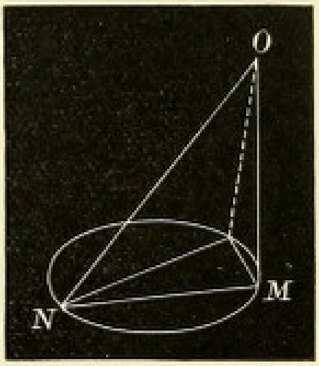


Fig. 172.

[1] Stokes, *Brit. Assoc. Report*, 1862.

vibrations are at right angles; they are therefore polarised at right angles. To verify this it is only necessary to receive the emergent cylinder of light on a tourmaline plate or a Nicol's prism, and of the two extremities of the same diameter of the ring one will be completely dark and the other brightest, the illumination gradually fading round the ring from the latter point to the former.

**203. Axes of Single Ray Velocity.**—We have seen that the wave surface presents four singular points in the plane $zx$. These points P, P′, $P_1$, $P'_1$, are common to both sheets of the surface, and are such that at any one of them P an infinite number of tangent planes can be drawn to the surface, and not merely two, as Fresnel appears to have imagined, viz.—one to the circle and one to the ellipse (Fig. 170). This system of tangent planes forms a tangent cone to the surface at the conical point P. Now if a ray travels in any direction in the crystal, the velocity of the ray is measured by the radius vector of the wave surface drawn in its direction. Consequently in any direction we have in general two ray velocities, since the radius vector has in general two values, one given by each sheet of the wave. But if a ray travels in the direction OP there is only one value of the radius vector, and consequently only one velocity of the ray. Consequently both rays travel in the direction OP (or OP′) with the same velocity, and these directions are called the *axes of single ray velocity*. They are generally very close to the optic axes, or axes of single wave velocity, but they are not on that account to be confounded with them.

The angle between the axes of single ray velocity may be easily expressed in terms of $a$, $b$, $c$. For the co-ordinates $x$ and $z$ of P are common to the circle

$$x^2+z^2=b^2,$$

and to the ellipse

$$a^2x^2+c^2z^2=a^2c^2,$$

therefore

$$x=c\sqrt{\frac{a^2-b^2}{a^2-c^2}}, \quad \text{and } z=a\sqrt{\frac{b^2-c^2}{a^2-c^2}}.$$

Hence if the angle PO$x$ (Fig. 170) be denoted by $\psi$, we have $\tan\psi=z/x$, and therefore

$$\cos\psi=\frac{c}{b}\sqrt{\frac{a^2-b^2}{a^2-c^2}}, \quad \sin\psi=\frac{a}{b}\sqrt{\frac{b^2-c^2}{a^2-c^2}}, \quad \tan\psi=\frac{a}{c}\sqrt{\frac{b^2-c^2}{a^2-b^2}}.$$

The right lines joining O to P and $P_1$ are consequently given by the equations

$$z=\pm\frac{a}{c}\sqrt{\frac{b^2-c^2}{a^2-b^2}}\cdot x, \quad \text{and } y=0;$$

they are therefore perpendicular to the circular sections of the reciprocal ellipsoid

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}+\frac{z^2}{c^2}=1.$$

*Cor.*—The above values of the co-ordinates of P show that it is a singular point on the surface, for they satisfy the equations

$$\frac{dS}{dx}=0, \quad \frac{dS}{dy}=0, \quad \frac{dS}{dz}=0,$$

where S denotes the equation of the wave surface. There is consequently at P a tangent cone to the surface.

The axes of single ray velocity are called the *axes of external conical refraction.*

**204. External Conical Refraction.**—The direction pursued by a refracted ray, after emerging from the crystal, is determined by the position of the tangent plane to the wave surface at the point where the ray meets it. But at any one of the conical points P there is an infinite number of tangent planes enveloping a cone, consequently the ray which traverses the crystal in the direction OP (or OP′) may on emergence pursue the direction determined by any one of these tangent planes. The emergent beam should therefore be of a conical form.

Dr. Lloyd found that this was fully verified by experiment. Taking the plate of aragonite already mentioned, he placed on each face of it a thin plate of metal, perforated by a very minute aperture, as shown in Fig. 173. These plates were so adjusted that the line connecting the two apertures coincided with the direction of the axis of single ray velocity. A flame of a lamp was then brought near the aperture O in such a manner that the central part of the convergent beam should have an incidence of 15° or 16°. When the adjustment was completed a brilliant annulus of light was seen on looking through the aperture P in the second plate. Whenever the second plate was ever so slightly moved, so that the line OP connecting the apertures no longer coincided with the axis of single ray velocity, the phenomenon rapidly changed and the annulus resolved itself into two separate images.

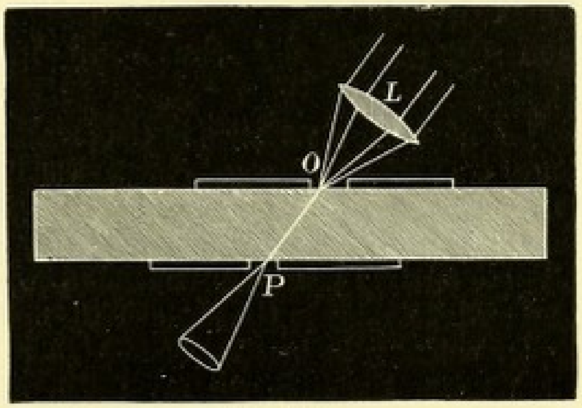


Fig. 173.

The incident light was also brought to a focus on the surface of the plate by means of a lens of short focal length. In this case the upper plate was dispensed with and the lamp removed to a distance. The rays of the sun were also used and the emergent light received on a screen. Each ray of the incident cone of light is doubly refracted at O, but of all the rays on a certain conical shell of the incident cone of light one of the refracted rays will travel along OP, and emerge as a ray of the external cone. If the aperture in the second plate is not very small, some of the rays which do not travel exactly along OP will be allowed through, and considerable discrepancy will occur between the results of observation and theory. However, when the necessary correction is applied, the agreement between the theoretical and observed magnitude of the angle of the cone was found to be nearly complete, the observed angle being 2° 59′ and the calculated angle 3° 0′ 58″.

Dr. Lloyd also determined that when external conical refraction is exhibited the ray OP is parallel to the axis of single ray velocity. In order to do this he observed the angle of incidence of the axis of the convergent beam on the first face, and he found it 15° 58′ when the conical refraction occurred. He also calculated by theory the angle at which the axis of the incident cone should meet the first face in order that the refracted ray should be parallel to the axis of single ray velocity, and he found it to be[1] 15° 25′ 8″.

It was also found by experiment that "*the angle between the planes of polarisation of any two rays of the cone is half the angle between the planes containing the rays themselves and the axis.*" This remarkable law is also in complete accordance with the theory as in the case of internal conical refraction (Art. 202).

**205. Relation between the Velocities of Propagation of a Plane Wave and the Position of the Wave Normal with respect to the Optic Axes.**—The velocities of the two waves travelling in any direction are given by the radii in that direction of the surface

[1] A remarkable variation of the phenomena took place on substituting a narrow *linear aperture* for the small circular one in the plate next the lamp, the line being so adjusted that the plane passing through it and the aperture on the second face should coincide with the plane of the optic axes. In this case, according to the hitherto received views, all the rays transmitted through the second aperture should be refracted doubly in the plane of the optic axes so that no part of the line should appear enlarged in breadth in looking through this aperture; while according to Sir William Hamilton the ray which proceeds in the direction OM should be refracted *in every plane*. The latter was found to be the case; in the neighbourhood of each of the optic axes the luminous line was bent, on either side of the plane of the axes, into an oval curve. This curve, it is easy to show, is the *conchoid of Nicomedes*, whose asymptote is the line on the first surface (Lloyd, *Wave Theory of Light*, p. 212).

of normal velocities; they are consequently determined by the equation

$$\frac{l^2}{r^2-a^2}+\frac{m^2}{r^2-b^2}+\frac{n^2}{r^2-c^2}=0,$$

where $l$, $m$, $n$ are the direction cosines of $r$. Hence

$$r^4-[(b^2+c^2)l^2+(c^2+a^2)m^2+(a^2+b^2)n^2]r^2+\Sigma b^2c^2l^2=0,$$

or, denoting the roots of this equation by $r'^2$ and $r''^2$, we have

$$r'^2+r''^2=\Sigma(b^2+c^2)l^2,$$

and

$$r'^2r''^2=\Sigma b^2c^2l^2.$$

Now if $r$ makes angles $\theta'$ and $\theta''$ with the optic axes, and, using the notation of Art. 201, if the direction angles of the optic axes be $\phi$, $\frac{1}{2}\pi$, $\frac{1}{2}\pi-\phi$, and $\pi-\phi$, $\frac{1}{2}\pi$, $\frac{1}{2}\pi-\phi$ respectively, we have

$$\cos\theta'=l\cos\phi+n\sin\phi,$$
$$\cos\theta''=-l\cos\phi+n\sin\phi.$$

Therefore

$$l=\frac{\cos\theta'-\cos\theta''}{2\cos\phi}=\tfrac{1}{2}(\cos\theta'-\cos\theta'')\sqrt{\frac{a^2-c^2}{a^2-b^2}},\qquad\text{(Art. 201)},$$

$$n=\frac{\cos\theta'+\cos\theta''}{2\sin\phi}=\tfrac{1}{2}(\cos\theta'+\cos\theta'')\sqrt{\frac{a^2-c^2}{b^2-c^2}},\qquad\text{,,}$$

and

$$m^2=1-l^2-n^2.$$

Hence

$$r'^2+r''^2=a^2+c^2-\tfrac{1}{4}(\cos\theta'-\cos\theta'')^2(a^2-c^2)+\tfrac{1}{4}(\cos\theta'+\cos\theta'')^2(a^2-c^2),$$
$$=a^2+c^2+(a^2-c^2)\cos\theta'\cos\theta'',$$

and

$$r'^2r''^2=a^2c^2+\tfrac{1}{4}(a^2-c^2)^2(\cos^2\theta'+\cos^2\theta'')+\tfrac{1}{2}(a^4-c^4)\cos\theta'\cos\theta''.$$

From which we have

$$(r'^2-r''^2)^2=(r'^2+r''^2)^2-4r'^2r''^2=(a^2-c^2)^2\sin^2\theta'\sin^2\theta'',$$

or finally, since $r'$ and $r''$ measure the velocities, we have

$$v'^2-v''^2=r'^2-r''^2=(a^2-c^2)\sin\theta'\sin\theta'',$$

and

$$v'^2+v''^2=a^2+c^2+(a^2-c^2)\cos\theta'\cos\theta'',$$

which establishes a relation between the velocities of the two plane waves which are propagated in any direction and the angles which this direction makes with the optic axes.

*Cor.*—The normal velocities of the two waves propagated in the same direction are

$$v'^2 = \tfrac{1}{2}(a^2+c^2) + \tfrac{1}{2}(a^2-c^2)\cos(\theta'-\theta'')$$
$$v''^2 = \tfrac{1}{2}(a^2+c^2) + \tfrac{1}{2}(a^2-c^2)\cos(\theta'+\theta'').$$

**206. Relation connecting the Ray Velocities in a given Direction, and the Angles made with the Axes of Single Ray Velocity.**—Writing the equation of the wave surface in the form

$$r^4\Sigma a^2l^2 - r^2\Sigma a^2(b^2+c^2)l^2 + a^2b^2c^2 = 0,$$

and denoting its roots by $r'^2$ and $r''^2$, we have

$$\frac{1}{r'^2}+\frac{1}{r''^2}=\Sigma\left(\frac{1}{b^2}+\frac{1}{c^2}\right)l^2=\frac{1}{a^2}+\frac{1}{c^2}+l^2\left(\frac{1}{b^2}-\frac{1}{a^2}\right)+n^2\left(\frac{1}{b^2}-\frac{1}{c^2}\right),$$

and

$$\frac{1}{r'^2r''^2}=\Sigma\frac{l^2}{b^2c^2}.$$

Now if the radius vector $r$ makes angles $\theta'$ and $\theta''$ with the axes of single ray velocity, and if the direction angles of these lines be $\psi$, $\frac{1}{2}\pi$, $\frac{1}{2}\pi-\psi$, and $\pi-\psi$, $\frac{1}{2}\pi$, $\frac{1}{2}\pi-\psi$, respectively, then

$$\cos\theta' = l\cos\psi + n\sin\psi,$$
$$\cos\theta'' = -l\cos\psi + n\sin\psi.$$

Therefore

$$l=\frac{\cos\theta'-\cos\theta''}{2\cos\psi}=\tfrac{1}{2}(\cos\theta'-\cos\theta'')\frac{b}{c}\sqrt{\frac{a^2-c^2}{a^2-b^2}},$$

and

$$n=\frac{\cos\theta'+\cos\theta''}{2\sin\psi}=\tfrac{1}{2}(\cos\theta'+\cos\theta'')\frac{b}{a}\sqrt{\frac{a^2-c^2}{b^2-c^2}},$$

while $m^2$ is determined by the equation

$$m^2 = 1 - l^2 - n^2.$$

Hence

$$\frac{1}{r'^2}+\frac{1}{r''^2}=\frac{1}{a^2}+\frac{1}{c^2}+\left(\frac{1}{a^2}-\frac{1}{c^2}\right)\cos\theta'\cos\theta'',$$

and

$$\frac{1}{r'^2r''^2}=\frac{1}{a^2c^2}+\frac{1}{4}\left(\frac{1}{a^2}-\frac{1}{c^2}\right)^2(\cos^2\theta'+\cos^2\theta'')+\frac{1}{2}\left(\frac{1}{a^4}-\frac{1}{c^4}\right)\cos\theta'\cos\theta''.$$

Whence

$$\frac{1}{r'^2}-\frac{1}{r''^2}=\left(\frac{1}{a^2}-\frac{1}{c^2}\right)\sin\theta'\sin\theta''.$$

The difference of the squares of the reciprocals of the ray velocities is consequently proportional to the product of the sines of the angles which the common direction of the rays makes with the axes of single ray velocity. Denoting the velocities of the rays by $v'$ and $v''$ we may write this relation in the form

$$v'^{-2} - v''^{-2} = (a^{-2} - c^{-2}) \sin\theta' \sin\theta''.$$

*Cor.* 1.—Since the velocities are inversely proportional to the refractive indices, we have

$$\mu'^2 - \mu''^2 = (\mu_1^2 - \mu_3^2) \sin\theta' \sin\theta'',$$

or approximately

$$\mu' - \mu'' = (\mu_1 - \mu_3) \sin\theta' \sin\theta''.$$

*Cor.* 2.—Since the relative retardation introduced by a plate of thickness $e$ is $e(\mu' - \mu'')$ for normal incidence, we have

$$\delta = e(\mu' - \mu'') = e(\mu_1 - \mu_3) \sin\theta' \sin\theta''.$$

*Cor.* 3.—The ray velocities are given by the equations

$$v'^{-2} = \tfrac{1}{2}(c^{-2} + a^{-2}) + \tfrac{1}{2}(c^{-2} - a^{-2}) \cos(\theta' + \theta''),$$
$$v''^{-2} = \tfrac{1}{2}(c^{-2} + a^{-2}) + \tfrac{1}{2}(c^{-2} - a^{-2}) \cos(\theta' - \theta'').$$

## BIAXAL CRYSTALS

### PRINCIPAL INDICES FOR SODIUM LIGHT

| | Least $\left(\frac{V}{a}\right)$. | Mean $\left(\frac{V}{b}\right)$. | Greatest $\left(\frac{V}{c}\right)$. | Temp. | Observer. |
|---|---|---|---|---|---|
| Aragonite . | 1·53013 | 1·68157 | 1·68589 | ... | Rudberg. |
| Borax . . | 1·4463 | 1·4682 | 1·4712 | 23° | Kohlrausch. |
| Mica . . | 1·5609 | 1·5941 | 1·5997 | 23° | ,, |
| Nitre . . | 1·3346 | 1·5056 | 1·5064 | 16° | Schrauf. |
| Selenite . . | 1·52082 | 1·52287 | 1·53048 | 17° | V. Lang. |
| Sulphur (prism) | 1·9505 | 2·0383 | 2·2405 | 16° | Schrauf. |
| Topaz . . | 1·61161 | 1·61375 | 1·62109 | ... | Rudberg. |

# CHAPTER XIII

## REFLECTION AND REFRACTION OF POLARISED LIGHT

**207. Fundamental Principles and Hypotheses.** — The first attempt to determine the relation between the intensities of the incident, reflected, and refracted pencils when light falls upon the surface of separation of transparent media was made by Young,[1] but he confined his investigations to the particular case of perpendicular incidence. The amplitude, form, and phase of the incident vibration being given, the problem before us is to determine the amplitudes, forms, and phases of the reflected and refracted vibrations. Thus if the simple vibration

$$y = a \sin(\omega t + \alpha)$$

gives birth to the simple reflected and refracted vibrations

$$y_1 = b \sin(\omega t + \beta), \quad \text{and } y_2 = c \sin(\omega t + \gamma),$$

which differ from the incident in phase and amplitude, we require to determine $b$ and $c$, $\beta$ and $\gamma$ in terms of the known quantities.

The simplest case will be that in which there is no change of phase introduced by the reflection and refraction. The vibrations may then be written in the form

$$y = a \sin \omega t, \quad y_1 = b \sin \omega t, \quad y_2 = c \sin \omega t,$$

and the problem is reduced to the determination of $b$ and $c$. This case has been treated by Fresnel, but it is of very limited application, for in general the reflected and refracted vibrations differ in phase from the incident and from each other.

In approaching this subject hypotheses must be made in two departments, one with respect to the nature of the vibration, and the other with respect to the nature of the difference in the properties

[1] Art. "Chromatics," *Ency. Brit. Supplement.*

of the ether in different media, and as to whether the change of condition is sudden or gradual at their surface of separation.

The hypotheses adopted by Fresnel lead to formulæ which are in very close agreement with the results of experiment. He founded his theory (in 1821-23) on the following principles.[1]

(1) *The Principle of the Conservation of Energy*, from which it follows that the energy of the incident wave is equal to the sum of the energies of the reflected and refracted waves which arise from it. Denoting the amplitudes of the corresponding vibrations by $a$, $b$, $c$ respectively, we have the energy equation (Art. 68)

$$\rho(a^2 - b^2)\sin 2i = \rho' c^2 \sin 2r \qquad \text{(energy equation)},$$

connecting the amplitudes of the incident, reflected, and refracted vibrations.

This equation of course is deduced on the supposition that the second medium absorbs no part of the refracted pencil. In general, however, the second medium absorbs a portion of the refracted light, and the corresponding energy appears as heat in the elevation of temperature of the substance. The calculations and formulæ founded on this equation are consequently limited by this supposition, and they therefore apply only to the case of waves which are transmitted without absorption. It is further assumed that the transverse vibrations of the incident light excite only transverse (or light) vibrations in the second medium, or that the entire energy of the incident light appears again as light in the reflected and refracted pencils. In the case of elastic solids, however, both longitudinal and transverse vibrations are produced by reflection and refraction, so that we have in general two reflected and two refracted waves, one transverse and the other longitudinal, and these are propagated with different velocities.

(2) *Hypothesis of Uniform Elasticity of the Ether.*—Some hypothesis must now be made concerning the symbols $\rho$ and $\rho'$ which are called the densities of the ether in the two media. Fresnel assumes that the velocity of propagation in any medium varies inversely as the square root of the ether density in that medium, so that

$$\frac{\sqrt{\rho'}}{\sqrt{\rho}} = \frac{v}{v'} = \frac{\sin i}{\sin r}.$$

Now the velocity of propagation of waves in elastic matter is measured by the square root of the elasticity divided by the density; this assumption is consequently analogous to saying that the elasticity

[1] Fresnel, *Œuvres*, tom. i. pp. 441-799.

of the ether, or that property of it which corresponds to the elasticity of ordinary matter, is the same in all media. Introducing this assumption into the energy equation we have Fresnel's modified form

$$\frac{a^2-b^2}{c^2}=\frac{\tan i}{\tan r} \qquad \text{(Fresnel's energy equation).}$$

(3) *Continuity of the Displacement.*—To determine the ratios of $b$ and $c$ to $a$ we require another equation. This is obtained by supposing that the displacement remains the same in crossing the surface of separation. Thus if planes be drawn parallel to the interface and very near it, one in each medium, the velocities and displacements of the elements in these planes can only differ by an infinitely small fraction of their own value. If the ethers in the two media be treated as two portions of different elastic substances (like jellies for example) in contact, then at the interface they must always remain in contact—that is, during the motion there is no slipping of one on the other parallel to the surface, and also there should be no separation or relative motion perpendicular to the surface. The displacement at the common surface must be the same in the two media, and this must include the longitudinal displacement, or pressural wave, as well as the transverse vibrations which are supposed to constitute light.

According to the elastic-solid theory the ether belonging to any medium always remains in that medium, never crossing the interface or changing its density. If, however, we look upon the ether in the two media as being continuous but differing in density, a portion of the ether in either may cross the interface into the other, and a thin layer of the ether at the surface might suffer rapid periodic changes of density. However, if we admit Fresnel's assumption that there is no change of phase in crossing the surface, this layer of variable density must be infinitely thin compared with the length of a wave, so thin, in fact, that the phases of the vibrations on each side of it may be considered the same.[1]

Fresnel did not consider the component of the displacement perpendicular to the surface, he merely secures continuity parallel to the interface, so that there is no tangential slipping of the ether in one medium on that in the other.

MacCullagh, on the other hand, worked on the supposition that the vibrations in the two contiguous media are *equivalent*—that is, the resultant of the incident and reflected vibrations is the same, both in

[1] See further Glazebrook's "Report on Optical Theories," *Brit. Assoc. Report*, 1885, p. 186.

magnitude and direction, as the resultant refracted vibration. This hypothesis he termed the principle of *equivalent vibrations*.[1]

**208. Light Polarised in the Plane of Incidence.**—According to the theory of Fresnel the direction of the vibration is perpendicular to the plane of polarisation, so that for light polarised in the plane of incidence the vibration is parallel to the surface of separation. Again, since no change of phase is supposed to accompany the reflection and refraction, the extreme displacements $a$, $b$, $c$, will be attained at the same instant. Hence the complete displacement in the first medium is their algebraic sum $a+b$ (where $b$ may be inherently positive or negative with respect to $a$), and by the third principle at the interface this must be equal to $c$, the displacement in the second. Hence to determine $b$ and $c$ we have the equation of continuity

$$a+b=c \qquad (1),$$

together with Fresnel's energy equation,

$$a^2-b^2=c^2 \tan i \cot r \qquad (2).$$

Dividing (2) by (1) we obtain

$$a-b=c \tan i \cot r \qquad (3).$$

Combining (1) and (3) we find at once

$$b=-a\frac{\sin(i-r)}{\sin(i+r)} \qquad \text{(reflected)},$$

and

$$c=\frac{2a\cos i \sin r}{\sin(i+r)} \qquad \text{(refracted)}.$$

Thus the sign of $b$ is opposite to or the same as that of $a$ according as $i$ is greater or less than $r$—that is, according as the second medium is more or less refracting than the first.

It should be remarked that the relative intensities of the incident reflected and refracted rays are not measured merely by $a^2$, $b^2$, $c^2$, but by the rates at which energy is propagated by the corresponding waves (Art. 68)—that is, according to Fresnel's theory, by

$$a^2 : b^2 : c^2 \tan i \cot r,$$

or by

$$a^2 : a^2\frac{\sin^2(i-r)}{\sin^2(i+r)} : a^2\frac{\sin 2i \sin 2r}{\sin^2(i+r)}.$$

[1] MacCullagh, "On the Laws of Crystalline Reflection and Refraction," *Trans. Roy. Irish Acad.* vol. xviii. January 1837.

*Cor.* 1.—The expression for $b$ may be written in the form

$$b=-a\frac{\sin(i-r)}{\sin(i+r)}=-a\frac{\mu\cos r-\cos i}{\mu\cos r+\cos i}=-a\frac{\mu-\cos i/\cos r}{\mu+\cos i/\cos r}.$$

Hence if $i=r=0^\circ$ we have $\cos i=\cos r=1$, and therefore at perpendicular incidence

$$b=-a\frac{\mu-1}{\mu+1},$$

which is the expression arrived at by Young.

As $i$ increases from $0^\circ$ to $90^\circ$ the value of $b$ increases numerically from Young's value to $-a$, for

$$\frac{\cos i}{\cos r}=\left\{\frac{1-\sin^2 i}{1-(\sin^2 i)/\mu^2}\right\}^{\frac{1}{2}},$$

which diminishes continuously from unity to zero as $i$ increases from $0^\circ$ to $90^\circ$.

Similarly

$$c=\frac{2a\cos i\sin r}{\sin(i+r)}=\frac{2a\cos i}{\mu\cos r+\cos i}.$$

Therefore at perpendicular incidence

$$c=\frac{2a}{\mu+1},$$

consequently $c$ decreases from this value to zero as $i$ increases from zero to $\frac{1}{2}\pi$. The intensities at perpendicular incidence are proportional to

$$a^2,\quad a^2\left(\frac{\mu-1}{\mu+1}\right)^2,\quad a^2\frac{4\mu}{(\mu+1)^2}.$$

These formulæ have been verified photometrically by Arago, while Provostaye and Desains, by means of the thermopile, have confirmed their accuracy for heat radiations.

**209. Light Polarised Perpendicularly to the Plane of Incidence.**—In the case of light polarised perpendicularly to the plane of incidence the vibration, according to Fresnel's theory, is in the plane of incidence; but the vibration must be in the wave front, it is therefore along the direction AB (Fig. 46) and makes an angle $i$ with the surface of separation. In the reflected and refracted waves it is along A′B′ and A′C′, making angles $i$ and $r$ respectively with the surface. Hence the algebraic sum of the displacements in the upper medium is $(a+b)\cos i$ parallel to the surface, and in the lower medium $c\cos r$, consequently for no slipping at the surface of separation we have

$$(a+b)\cos i=c\cos r.$$

Combining this with the energy equation

$$a^2 - b^2 = c^2 \tan i \cot r$$

we have, by division,

$$a - b = c \sin i / \sin r.$$

Therefore

$$b = -a \frac{\tan (i-r)}{\tan (i+r)} \qquad \text{(reflected)},$$

and

$$c = \frac{2a \cos i \sin r}{\sin (i+r) \cos (i-r)} \qquad \text{(refracted)}.$$

Hence we see that if $i$ be greater than $r$ then $a$ and $b$ will have opposite signs; but, on the other hand, when the first medium is more refracting than the second, $a$ and $b$ have the same sign. The relative intensities are in the ratios $a^2 : b^2 : c^2 \tan i \cot r$—that is,

$$a^2, \quad a^2 \frac{\tan^2 (i-r)}{\tan^2 (i+r)}, \quad a^2 \frac{\sin 2i \sin 2r}{\sin^2 (i+r) \cos^2 (i-r)}.$$

*Cor.* 1.—Writing the expression for $b$ in the form

$$b = -a \cdot \frac{\mu \cos r - \cos i}{\mu \cos r + \cos i} \cdot \frac{\cos (i+r)}{\cos (i-r)},$$

obtained from the above value by merely dividing the numerator and denominator by $\sin r$, we see that when $i = r = 0$ we have

$$b = -a \frac{\mu - 1}{\mu + 1},$$

the square of which measures the intensity of the light reflected normally. This expression is the same as that which determined the normally reflected beam when the incident light is polarised in the plane of incidence, as it obviously should be, for in both cases when the light is incident normally, the vibration is parallel to the surface, and the two should be reflected according to the same law.

Similarly when $i = r = 0$ the expression for $c$ becomes the same in both cases, and the reflected and refracted intensities are complementary.

*Cor.* 2.—If $i + r = 90°$, which is the general condition at the angle of maximum polarisation, we have

$$b = 0, \quad \text{and } c = a/\mu.$$

The light is therefore all refracted, and its intensity, measured by $c^2 \tan i \cot r$, becomes at once $c^2 \mu^2$ (since $i + r = 90°$), or $a^2$, which is the measure of the intensity of the incident light.

Hence the amplitude of the reflected pencil decreases from Young's value to zero, as the angle of incidence increases from zero to the polarising angle ($\tan i = \mu$). It then changes sign (or the phase changes by $\pi$) as $i$ increases, and attains the value $a$ at grazing incidence.

**210. Light Polarised in any Plane—Rotation of the Plane of Polarisation by Reflection.**—Let us now suppose the incident light to be plane-polarised in a plane inclined at any angle $\alpha$ to the plane of incidence. The direction of the vibration now makes an angle $90° - \alpha$ with the plane of incidence (following Fresnel), and we can therefore resolve it into two components

$$a \cos \alpha, \quad \text{and } a \sin \alpha,$$

perpendicular to, and in, the plane of incidence respectively. The former component may be regarded as a beam of light polarised in the plane of incidence, and it will give rise to reflected and refracted rays determined by Art. 208, in which $a$ is to be replaced by $a \cos \alpha$. The latter component is a pencil polarised perpendicularly to the plane of incidence, and also gives rise to reflected and refracted rays determined by Art. 209.

Hence the reflected light is the resultant of two portions, one polarised in the plane of incidence and the other polarised perpendicularly to it, the amplitudes of these portions are

$$-a \cos \alpha \frac{\sin(i-r)}{\sin(i+r)}, \quad \text{and} \quad -a \sin \alpha \frac{\tan(i-r)}{\tan(i+r)} \qquad \text{(reflected)}$$

respectively, while the refracted light consists of

$$2a \cos \alpha \frac{\cos i \sin r}{\sin(i+r)}, \quad \text{and} \quad 2a \sin \alpha \frac{\cos i \sin r}{\sin(i+r)\cos(i-r)} \qquad \text{(refracted)}$$

polarised in and perpendicular to the plane of incidence respectively. If then the reflected light be plane-polarised in a plane making an angle $\beta$ with the plane of reflection, its components perpendicular to and in this plane are $b \cos \beta$ and $b \sin \beta$ respectively, where $b$ is the amplitude of the reflected vibration. So also if $c$ be the amplitude of the refracted vibration, and $\gamma$ the angle its plane of polarisation makes with the plane of refraction, its components perpendicular to and in that plane are $c \cos \gamma$ and $c \sin \gamma$ respectively, consequently we have

$$\left.\begin{aligned} b \cos \beta &= -a \cos \alpha \frac{\sin(i-r)}{\sin(i+r)} \\ b \sin \beta &= -a \sin \alpha \frac{\tan(i-r)}{\tan(i+r)} \end{aligned}\right\} \qquad \text{(reflected)},$$

$$\left.\begin{aligned} c \cos \gamma &= 2a \cos \alpha \frac{\cos i \sin r}{\sin(i+r)} \\ c \sin \gamma &= 2a \sin \alpha \frac{\cos i \sin r}{\sin(i+r)\cos(i-r)} \end{aligned}\right\} \qquad \text{(refracted)}.$$

From these equations we can determine the positions of the planes of polarisation of the reflected and refracted pencils, for we have at once

$$\tan\beta = \tan\alpha \frac{\cos(i+r)}{\cos(i-r)},$$

and

$$\tan\gamma = \tan\alpha \sec(i-r).$$

Hence[1]

$$\tan\beta = \tan\gamma \cos(i+r).$$

So also for the amplitudes of the vibrations, we have

$$b^2 = a^2 \left\{ \sin^2\alpha \frac{\tan^2(i-r)}{\tan^2(i+r)} + \cos^2\alpha \frac{\sin^2(i-r)}{\sin^2(i+r)} \right\}$$

and

$$c^2 = \frac{4a^2 \cos^2 i \sin^2 r}{\sin^2(i+r)} \left\{ \frac{\sin^2\alpha}{\cos^2(i-r)} + \cos^2\alpha \right\},$$

$$= \frac{4a^2 \cos^2 i \sin^2 r}{\sin^2(i+r)} \left\{ 1 + \sin^2\alpha \tan^2(i-r) \right\},$$

while the intensities of the incident, reflected, and refracted beams are respectively proportional to

$$a^2, \quad b^2, \quad c^2 \tan i \cot r.$$

*Cor.* 1.—If $i + r = 90^\circ$, we have $\beta = 0$—that is, if the light is incident at the angle of maximum polarisation, the reflected light is polarised in the plane of reflection.

*Cor.* 2.—From the expression for $\tan\beta$, it is clear that while $i + r$ is less than $90^\circ$, we have $\cos(i+r) < \cos(i-r)$, and therefore $\tan\beta < \tan\alpha$, while as $i$ increases ($\alpha$ remaining constant) $\beta$ passes through zero as $i + r$ passes through $90^\circ$, and becomes negative when $i + r$ increases beyond $90^\circ$. The numerical value of $\tan\beta$ is therefore always less than that of $\tan\alpha$, or the effect of the reflection is to bring the plane of polarisation nearer to the plane of incidence, and at the angle of maximum polarisation the two coincide.

*Cor.* 3.—If the first medium is more refracting than the second, total reflection will occur when $r = 90^\circ$, and in this case $\tan\beta$ will be equal to $-\tan\alpha$, so that $\alpha$ and $\beta$ are supplementary.

*Cor.* 4.—If the same pencil of plane-polarised light be reflected $n$

[1] The relation connecting $\alpha$, $\beta$, and $\gamma$ may be stated in the form—"The vibrations in the incident and reflected waves coincide with the projections of the refracted vibration on those waves, or, the planes of vibration of the three waves intersect in a line which is the direction of the refracted vibration" (MacCullagh, *Trans. Roy. Irish Academy*, vol. xviii. 1837).

times at the same incidence $i$, and in the same plane we have for the azimuth of the pencil after the $n$ reflections

$$\tan \beta_n = \tan \alpha \frac{\cos^n (i+r)}{\cos^n (i-r)},$$

while if the refracted beam suffers $n$ refractions by passing through parallel plates of the two media, we have

$$\tan \gamma_n = \tan \alpha \sec^n (i-r).$$

For by refraction into a plate we have $\tan \gamma_1 = \tan \alpha \sec (i-r)$, and by the refraction out at the second surface we have $\tan \gamma_2 = \tan \gamma_1 \sec (i-r)$, or

$$\tan \gamma_2 = \tan \alpha \sec^2 (i-r).$$

In passing through $n$ parallel plates the ray suffers $2n$ refractions, and

$$\tan \gamma = \tan \alpha \sec^{2n} (i-r).$$

The general effect of reflection or refraction is therefore a rotation of the plane of polarisation.

**211. Elliptically and Circularly Polarised Light.**—We have seen already (Art. 47) that two rectangular vibrations differing in amplitude and phase compound into an elliptic vibration, and conversely we may decompose an elliptic vibration into two rectangular vibrations of the form

$$x = a \sin \omega t, \quad \text{and } y = b \sin (\omega t + \delta).$$

Consequently, if the incident light be elliptically polarised, we may resolve it into two components, one polarised in the plane of incidence and the other polarised perpendicularly to it, and apply the foregoing formulæ to determine the nature and intensity of the reflected and refracted rays. The general expressions for these elliptic vibrations can be formed without difficulty, but as their discussion presents no particular interest, we shall only consider the case of circularly polarised light.

If the incident light be circularly polarised, its component vibrations may be represented by

$$x = a \cos \omega t, \quad \text{and } y = a \sin \omega t.$$

The components of the reflected and refracted vibrations may be written down from the formulæ of Arts. 208, 209; thus for the reflected vibration we have

$$x = -a \frac{\sin (i-r)}{\sin (i+r)} \cos \omega t = -\text{A} \cos \omega t,$$

$$y = -a \frac{\tan (i-r)}{\tan (i+r)} \sin \omega t = -\text{B} \sin \omega t,$$

so that

$$\frac{x^2}{A^2}+\frac{y^2}{B^2}=1,$$

or the reflected light is elliptically polarised, the axes of the ellipse being respectively in, and perpendicular to, the plane of reflection.

If the light is incident perpendicularly, we have $i=r=0$ and

$$A=B=-a(\mu-1)/(\mu+1),$$

so that the reflected light is circularly polarised and its intensity is

$$a^2\left(\frac{\mu-1}{\mu+1}\right)^2.$$

It is to be remarked, however, that the sense of the circular vibration in the reflected light is opposite to that in the incident light, for the $x$ component of the vibration is affected with the negative sign. This theoretical deduction of Earnshaw has been verified experimentally by Powell.[1]

The character of the refracted light may be investigated in a similar manner. It is to be observed in all cases that the $y$ component of the reflected vibration vanishes when $i+r=90°$, or the reflected light is always plane-polarised in the plane of reflection at the angle of maximum polarisation.

**212. Reflection and Refraction of Common Light.**—The characteristic of ordinary light is that on transmission through a doubly refracting crystal, such as Iceland spar, it is divided into two pencils of equal intensity. If the crystal be also doubly absorbing, the ordinary and extraordinary rays will be of unequal intensity, and (as in the case of tourmaline) only one of them may be transmitted. On entering the crystal, however, the light is divided into two equal beams vibrating in two rectangular planes, and this whatever be the orientation of the crystal. Hence if ordinary light be resolved into any two rectangular components they will be equal, and therefore the components of the incident beam in and at right angles to the plane of incidence will be equal in intensity, and each half that of the incident light. Consequently the intensity of the reflected light will be (Arts. 208, 209)

$$I=\tfrac{1}{2}a^2\left[\frac{\sin^2(i-r)}{\sin^2(i+r)}+\frac{\tan^2(i-r)}{\tan^2(i+r)}\right],$$

and the intensity of the refracted light will be

$$I'=\tfrac{1}{2}a^2\left[\frac{\sin 2i\sin 2r}{\sin^2(i+r)}+\frac{\sin 2i\sin 2r}{\sin^2(i+r)\cos^2(i-r)}\right].$$

[1] Powell, *Phil. Mag.* (3), vol. xxii. pp. 92, 262.

In the case of the reflected light the second term within the bracket represents the light polarised perpendicularly to the plane of incidence. If the two terms were equal, then the reflected light would be like the common light of the incident beam. The second term, however, is always less than the first, except under nearly grazing incidence, consequently there will in general be an excess of light polarised in the plane of incidence—that is, the reflected light will be partially polarised.

At the particular incidence $i + r = \frac{1}{2}\pi$ the second term vanishes entirely, and the reflected light is wholly polarised in the plane of incidence.

Again the first term in the expression for the intensity of the refracted light is less than the second, and this indicates that there is an excess of the refracted light polarised perpendicularly to the plane of incidence, but the whole refracted light is not completely polarised in this plane at the polarising angle. At this incidence the reflected light alone is completely plane-polarised, and the two beams contain *equal quantities* of polarised light.

**213. Total Reflection.**—When the first medium is more highly refracting than the second, total reflection occurs when the angle of incidence reaches a certain critical value determined by $\mu \sin i = 1$. The corresponding value of $r$ is $90^\circ$, and if $i$ be increased beyond its critical value a value of $\sin r$ greater than unity is required, so that the angle $r$ becomes imaginary, and the formulæ which determine the amplitudes and intensities of the reflected and refracted pencils are no longer applicable.

Experiment proves that when $i$ exceeds its critical value the refracted ray ceases to exist, and the reflected ray is equal in intensity to the incident. Theory confirms this result, for the construction of Huygens shows that in the second medium no real wave envelope can be drawn, and the conservation of energy then requires that the intensity of the reflected ray should be equal to that of the incident.

Taking the case of light polarised in the plane of incidence, we have for the amplitude of the reflected ray

$$b = -a\frac{\sin(i-r)}{\sin(i+r)} = a\frac{\mu \sin i \cos i - \sin i\sqrt{1-\mu^2 \sin^2 i}}{\mu \sin i \cos i + \sin i\sqrt{1-\mu^2 \sin^2 i}},$$

by writing $\sin r = \mu \sin i$, $\mu$ being the index of refraction from the less to the more refracting medium. Hence when $i$ increases beyond the critical value we have $\mu \sin i > 1$, and the expression for $b$ becomes imaginary. This imaginary form has been interpreted by Fresnel in the following manner, which is undoubtedly ingenious, but which must be regarded rather as an interesting curiosity than as a rigorous demonstration.

Multiplying the numerator and denominator of the expression for $b$ by the conjugate form of the denominator, it becomes

$$a\frac{\mu^2\sin^2 i\cos^2 i+\sin^2 i(1-\mu^2\sin^2 i)-2\mu\sin^2 i\cos i\sqrt{1-\mu^2\sin^2 i}}{\mu^2\sin^2 i\cos^2 i-\sin^2 i(1-\mu^2\sin^2 i)}.$$

The denominator of this expression is simply $(\mu^2-1)\sin^2 i$, so that on dividing above and below by $\sin^2 i$ it reduces to

$$a\left\{\frac{\mu^2+1-2\mu^2\sin^2 i}{\mu^2-1}-\frac{2\mu\cos i}{\mu^2-1}\sqrt{\mu^2\sin^2 i-1}\sqrt{-1}\right\}=a(\mathrm{P}-\mathrm{Q}\sqrt{-1})$$

where

$$\mathrm{P}=\frac{\mu^2+1-2\mu^2\sin^2 i}{\mu^2-1},\quad \text{and } \mathrm{Q}=\frac{2\mu\cos i}{\mu^2-1}\sqrt{\mu^2\sin^2 i-1},$$

and therefore

$$\mathrm{P}^2+\mathrm{Q}^2=1.$$

Hence if we take $\mathrm{P}=\cos\delta$ we will have $\mathrm{Q}=\sin\delta$, and the expression for $b$ becomes

$$b=a(\cos\delta-\sqrt{-1}\sin\delta).$$

Hence if the equation of the incident vibration be $y=a\sin\omega t$, the equation of the reflected vibration will be

$$y=a(\cos\delta-\sqrt{-1}\sin\delta)\sin\omega t.$$

Fresnel suspected that since the occurrence of $\sqrt{-1}$ in geometry indicates a rotation of $90^\circ$ in the position of the line whose length is multiplied by it, so it is probable that here the imaginary quantity $\sqrt{-1}$ denotes a change of $\frac{1}{2}\pi$ in the phase of the vibration to which it is attached. Proceeding on this assumption the equation of the reflected vibration becomes

$$\begin{aligned} y&=a\{\cos\delta\sin\omega t-\sin\delta\sin(\omega t+90^\circ)\},\\ &=a(\cos\delta\sin\omega t-\sin\delta\cos\omega t),\\ &=a\sin(\omega t-\delta). \end{aligned}$$

The interpretation, therefore, according to Fresnel is that the phase of the reflected vibration has been altered at reflection by an amount $\delta$ given by the equation

$$\tan\delta=\frac{\mathrm{Q}}{\mathrm{P}}=\frac{2\mu\cos i\sqrt{\mu^2\sin^2 i-1}}{\mu^2+1-2\mu^2\sin^2 i},$$

while the amplitude of the reflected vibration is equal to that of the incident.

In the same manner we may treat the case of light polarised at right angles to the plane of incidence, and the corresponding quantities P′, Q′, δ′ will be found to be as follows :

$$P' = \frac{\mu^2+1-(\mu^4+1)\sin^2 i}{(\mu^2-1)\{(\mu^2+1)\sin^2 i - 1\}},$$

$$Q' = \frac{2\mu\cos i\sqrt{\mu^2\sin^2 i - 1}}{(\mu^2-1)\{(\mu^2+1)\sin^2 i - 1\}},$$

$$\tan\delta' = \frac{2\mu\cos i\sqrt{\mu^2\sin^2 i - 1}}{(\mu^2+1)-(\mu^4+1)\sin^2 i}.$$

When the light is polarised in any azimuth, we may resolve it into two components, one polarised in the plane of incidence and the other polarised perpendicularly to it, and when total reflection occurs the former will suffer a change of phase $\delta$ and the latter a change $\delta'$, determined as above. The difference $\delta - \delta'$ is all that concerns us experimentally, and we have

$$\cos(\delta - \delta') = \frac{1-(\mu^2+1)\sin^2 i + 2\mu^2\sin^4 i}{(\mu^2+1)\sin^2 i - 1},$$

the reflected beam should therefore be elliptically polarised, the phase difference of the two components being $\delta - \delta'$.

*Cor.* 1.—If $\delta - \delta' = 0$—that is, if the change of phase produced by the reflection is the same for the component in the plane of incidence as for that perpendicular to it, the reflected ray will be plane-polarised. In this case $\cos(\delta - \delta') = 1$, and we have, therefore,

$$1-(\mu^2+1)\sin^2 i + 2\mu^2\sin^4 i = (\mu^2+1)\sin^2 i - 1,$$

or

$$\mu^2\sin^4 i - (\mu^2+1)\sin^2 i + 1 = 0.$$

Hence

$$\sin^2 i = \frac{(\mu^2+1)\pm(\mu^2-1)}{2\mu^2},$$

$$= 1 \text{ or } 1/\mu^2.$$

The first value of $\sin i$ corresponds to grazing incidence, and the second to the limiting incidence for total reflection. At the two limits of total reflection, therefore, the reflected light is totally plane-polarised.

*Cor.* 2.—The difference of phase $\delta - \delta'$ passes through a maximum value at an angle of incidence determined by the equation

$$(\mu^2+1)\sin^2 i = 2,$$

and the corresponding maximum is given by

$$\cos(\delta - \delta') = \frac{8\mu^2 - (\mu^2+1)^2}{(\mu^2+1)^2} = \frac{4\mu^2 - (\mu^2-1)^2}{(\mu^2+1)^2}.$$

*Cor.* 3.—If $\cos(\delta - \delta') = 0$, the axes of the reflected elliptic vibration are in, and perpendicular to, the plane of incidence, and the corresponding incidence is given by

$$2\mu^2 \sin^4 i - (\mu^2 + 1) \sin^2 i + 1 = 0.$$

If at this incidence the plane of polarisation of the incident light makes an angle of 45° with the plane of incidence, the reflected light should be circularly polarised. This affords an experimental means of verifying the foregoing theoretical results, but in order that the values of $i$, determined by the foregoing quadratic, should be real, it is necessary to have

$$\mu^2 > 3 + \sqrt{8},$$

consequently, a substance of refractive index nearly equal to 3 would be required.

The required difference of phase may however be produced by two or more total reflections from a substance such as glass with a smaller index of refraction. Thus if the azimuth of the plane of polarisation of the incident light be 45°, and if a difference of phase of 45° be introduced at each reflection, the light twice reflected will be circularly polarised, or if a difference of phase of $\pi/2n$ be produced at each reflection, the light $n$ times totally reflected will be circularly polarised. In the first case the angle of incidence is determined by the equation

$$\cos(\delta - \delta') = 1/\sqrt{2},$$

or

$$4\mu^2 \sin^4 i - (2 + \sqrt{2})(\mu^2 + 1) \sin^2 i + 2 + \sqrt{2} = 0.$$

This gives real values of $i$ for glass and media whose index of refraction lies between 1·4 and 1·6.

We may at once write down the corresponding formulæ for three, or four, total reflections —that is, for $\delta - \delta' = \pi/6$, $\pi/8$, etc.

**214. Fresnel's Rhomb.**—In verification of the foregoing conclusions Fresnel constructed a parallelopiped of glass such that a ray of light AB (Fig. 174) falling normally on the end suffers total reflection internally at B, where it falls upon the face at an incidence of 55°, and again at C, and then emerges normally at the other end of the rhomb. At B a difference of phase of 45° is introduced—that is, $\frac{1}{8}\lambda$ retardation—and the same difference is again produced at C, so that in all 90° difference of phase is intro-

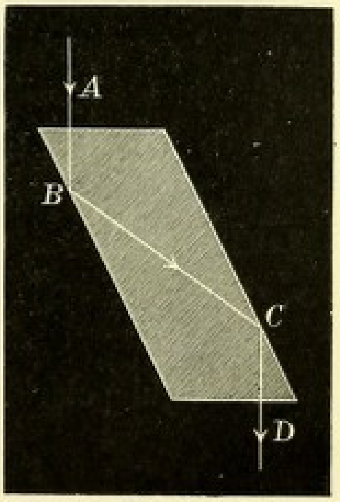


Fig. 174.

duced, or a retardation of $\frac{1}{4}\lambda$, and thus if the incident light be polarised at an angle of 45° to the plane of incidence ABC, its components in and perpendicular to the plane of incidence will be equal, and the emergent light is found to be circularly polarised.

Conversely, if the incident light be circularly polarised, the rhomb introduces a further difference of phase of 90°, so that the emergent light is plane-polarised. Hence, generally, if a ray of light originally plane-polarised in an azimuth of 45° to the plane of incidence be passed through any number ($n$) of Fresnel's rhombs, the emergent light will be circularly or plane-polarised according as $n$ is odd or even. With such rhombs we may therefore test between ordinary light and circularly polarised light.

By means of this rhomb we may also convert elliptically polarised light into plane-polarised light. For if the axes of the elliptic vibration be in and perpendicular to the plane of incidence ABC, the two internal reflections will introduce a further difference of phase of 90° between its components, and the emergent light should be plane-polarised. This light will then be extinguished by a Nicol's prism, and we can therefore test between elliptically polarised light and partially polarised light.

Fresnel's rhomb consequently possesses all the properties of a quarter-wave plate (Art. 227), and is preferable to it in working with white light, since the difference of the velocities of the different colours in glass is inconsiderable. However, the emergent beam of light changes its direction when the rhomb is rotated, so that there is trouble in following it with the other parts of the apparatus. It is best therefore to use a quarter-wave plate in working with monochromatic light.

**215. Newton's Rings—Polarisation of the Light.**—In calculating the effects produced by the reflection of light from very thin plates of transparent material (Art. 113) we supposed the amplitudes of the incident, reflected, and refracted vibrations to be $a$, $ab$, $ac$ respectively. We now know what the quantities $b$ and $c$ are in the case of polarised light. Thus for light polarised in the plane of incidence, taking $a=1$, we have

$$b=-\frac{\sin(i-r)}{\sin(i+r)}, \quad c=\frac{2\cos i\sin r}{\sin(i+r)},$$

and for light polarised in the perpendicular plane

$$b'=-\frac{\tan(i-r)}{\tan(i+r)}, \quad c'=\frac{2\cos i\sin r}{\sin(i+r)\cos(i-r)}.$$

Substituting this value of $b$ (or $b'$) in the expression of Art. 113, we obtain the intensity of the reflected light.

It is remarkable that when $i+r=90°$, we have $b'=0$, and consequently the reflected rings should disappear completely, and the transmitted light should be equal to the incident, when the light is polarised perpendicularly to the plane of incidence and incident at the polarising angle.

Again $b'$, after vanishing at the polarising angle, changes sign as the incidence increases beyond that angle. Consequently if the rings begin from a dark centre with $i+r<90°$, they should begin from a white centre with $i+r>90°$. These results are verified by experiment. The rings produced by polarised light beginning from a dark centre vanish when the angle of incidence reaches the polarising angle, and reappear again encircling a white centre when the angle of incidence exceeds the polarising angle, if the incident light be polarised perpendicularly to the plane of incidence.

If the incident light is polarised in any azimuth $a$, we may, as before, resolve it into two components, one polarised in the plane of incidence and the other perpendicular to it. As these components on reflection are at right angles and do not differ in phase, the intensity of the reflected light is merely the sum of their intensities, or (Art. 113)

$$I=\cos^2 a\frac{4a^2b^2\sin^2\frac{1}{2}\delta}{(1-b^2)^2+4b^2\sin^2\frac{1}{2}\delta}+\sin^2 a\frac{4a^2b'^2\sin^2\frac{1}{2}\delta}{(1-b'^2)^2+4b'^2\sin^2\frac{1}{2}\delta},$$

with a similar expression for the transmitted rings, where $b$ and $b'$ are to be replaced by their values given above. Maxima and minima occur therefore in the case of light polarised in any plane, as in the case of light polarised in or perpendicular to the plane of incidence. The minima of the reflected rings are still zero, and the maxima equal to

$$\frac{4a^2b^2\cos^2 a}{(1+b^2)^2}+\frac{4a^2b'^2\sin^2 a}{(1+b'^2)^2}.$$

For the transmitted rings the maxima are equal to $a^2$, that is the same as the incident light, and the minima are equal to

$$a^2\frac{(1-b^2)^2\cos^2 a}{(1+b^2)^2}+a^2\frac{(1-b'^2)^2\sin^2 a}{(1+b'^2)^2},$$

while there is no incidence for which the rings are black.

In Art. 113 the reflected vibration has been written in the form $X\sin\omega t+Y\cos\omega t$, and its phase retardation is therefore determined by the ratio of the coefficients of $\cos\omega t$ and $\sin\omega t$. Denoting this phase retardation by $\delta_1$ we have in the case of light polarised in the plane of incidence

$$\tan\delta_1=\frac{Y}{X}=-\frac{1-b^2}{1+b^2}\cot\tfrac{1}{2}\delta=-\frac{2\cot\frac{1}{2}\delta}{\mu\dfrac{\cos r}{\cos i}+\dfrac{\cos i}{\mu\cos r}},$$

and for light polarised at right angles to the plane of incidence

$$\tan\delta'_1 = -\frac{1-b'^2}{1+b'^2}\cot\tfrac{1}{2}\delta = -\frac{2\cot\frac{1}{2}\delta}{\mu\dfrac{\cos i}{\cos r}+\dfrac{\cos r}{\mu\cos i}}.$$

If the incident light be polarised in any azimuth $a$, the difference of phase of the components reflected in, and perpendicular to, the plane of incidence respectively will be given by

$$\tan(\delta'_1-\delta_1)=\frac{\sin\delta\left(\mu-\dfrac{1}{\mu}\right)\left(\dfrac{\cos r}{\cos i}-\dfrac{\cos i}{\cos r}\right)}{4\cos^2\frac{1}{2}\delta+\left[\left(\mu+\dfrac{1}{\mu}\right)^2+\left(\dfrac{\cos r}{\cos i}-\dfrac{\cos i}{\cos r}\right)^2\right]\sin^2\frac{1}{2}\delta}.$$

The reflected light will therefore in general be elliptically polarised. The difference of phase will be zero when $\delta=n\pi$, where $n$ is any whole number, even or odd, that is, for the central lines of the dark rings and the bright rings respectively. At these parts of the rings the light will consequently be plane-polarised.

In the interval between two dark rings therefore the difference of phase rises from zero to a maximum and then falls again to zero at the middle of the interval. It then changes sign and passes through similar variations in the second half of the interval. Thus in the first quarter it rises from zero to a maximum, and in the second it falls again to zero. In the third quarter it changes from zero to a negative maximum, and in the remaining quarter it falls again to zero.

The maximum value of $\delta'_1-\delta_1$ is less than $90°$, and hence the reflected light cannot ever be circularly polarised, for the denominator of $\tan(\delta'_1-\delta_1)$ can never be zero.

Denoting as before the amplitudes of the reflected components, polarised in and perpendicular to the plane of incidence, by $b$ and $b'$, we have for the ratio of the intensities of the reflected components [1]

$$\frac{I}{I'}=\cot^2 a\,.\,\frac{b^2}{b'^2}\,.\,\frac{(1-b'^2)^2+4b'^2\sin^2\frac{1}{2}\delta}{(1-b^2)^2+4b^2\sin^2\frac{1}{2}\delta}.$$

For $\sin\frac{1}{2}\delta=0$ we have dark rings under all incidences, and for $\sin\frac{1}{2}\delta=1$ we have bright rings polarised in an azimuth $a_1$, where

$$\tan a_1=\tan a\frac{b'(1+b^2)}{b(1+b'^2)},$$

and if the principal section of the analyser be placed in this direction they disappear in the extraordinary image, and occupy the place of the dark transmitted rings.

[1] These formulæ and those for the transmitted system are deduced from the formulæ of Arts. 113 and 114.

The transmitted rings are complementary. The azimuth of the transmitted light is given by

$$\tan^2 \alpha_2 = \tan^2 \alpha \frac{(1-b'^2)^2}{(1-b^2)^2} \left\{ \frac{(1-b^2)^2 + 4b^2 \sin^2 \frac{1}{2}\delta}{(1-b'^2) + 4b'^2 \sin^2 \frac{1}{2}\delta} \right\}.$$

For the dark transmitted rings $\sin \frac{1}{2}\delta = 1$, and

$$\tan \alpha_2 = \tan \alpha \frac{(1-b'^2)(1+b^2)}{(1-b^2)(1+b'^2)} = \tan \alpha \frac{\dfrac{\mu \cos r}{\cos i} + \dfrac{\cos i}{\mu \cos r}}{\dfrac{\mu \cos i}{\cos r} + \dfrac{\cos r}{\mu \cos i}},$$

while for the bright transmitted rings $\sin \frac{1}{2}\delta = 0$, and

$$\tan \alpha_2 = \tan \alpha \frac{(1-b'^2)(1-b^2)}{(1-b^2)(1-b'^2)} = \tan \alpha,$$

or the plane of polarisation is not changed. In observing the rings through an analyser successively placed in the azimuths $\alpha$ and $\alpha_2$, two systems of rings are seen which correspond to those ordinarily seen by reflection and transmission.

**216. Theory of Neumann and MacCullagh.**—The principle of the continuity of the ether adopted by Fresnel demands that the displacements parallel to the surface of separation should be the same in both media. But it also requires the displacements perpendicular to the surface to be the same, consequently if we have the equation

$$(a+b) \cos i = c \cos r \quad \text{(no slipping)},$$

we should also have

$$(a-b) \sin i = c \sin r \quad \text{(no separation)}.$$

Hence by multiplication we obtain

$$(a^2 - b^2) \sin i \cos i = c^2 \sin r \cos r \qquad (1).$$

But by the conservation of energy we have

$$\rho(a^2 - b^2) \sin i \cos i = \rho' c^2 \sin r \cos r \qquad (2),$$

and therefore the equation derived from the principle of continuity will be in accordance with that derived from the conservation of energy if

$$\rho = \rho'.$$

Hence Fresnel's assumption that the density of the ether is different in different media is inconsistent with the continuity of the ether at right angles to the surface.

Adopting the above equations, Neumann and MacCullagh postulate that the density of the ether is the same in all media, but that its elasticity is different. On solution they find expressions for the amplitude of the reflected rays which differ from those arrived at by Fresnel only in that the expression for light vibrating in the plane of incidence is that which Fresnel arrived at for light vibrating perpendicularly to the plane of incidence, so that according to their theory the direction of the vibration lies in the plane of polarisation and not perpendicular to it as in the theory of Fresnel.

Thus for light vibrating in the plane of incidence, we have

$$(a+b)\cos i = c\cos r,$$
$$(a-b)\sin i = c\sin r.$$

Therefore

$$b = a\frac{\sin(i-r)}{\sin(i+r)} \qquad \text{(reflected)},$$

and

$$c = \frac{a\sin 2i}{\sin(i+r)} \qquad \text{(refracted)}.$$

While for light vibrating at right angles to the plane of incidence we have

$$a+b=c,$$

and

$$(a^2-b^2)\sin 2i = c^2\sin 2r.$$

Therefore

$$b = a\frac{\tan(i-r)}{\tan(i+r)} \qquad \text{(reflected)},$$

$$c = \frac{a\sin 2i}{\sin(i+r)\cos(i-r)} \qquad \text{(refracted)}.$$

The expressions for the refracted rays have been interchanged like those for the reflected, but they have also been altered in the ratio $\sin r : \sin i$.

As in Fresnel's theory, we have for the azimuths of the planes of polarisation of the reflected and refracted rays (Art. 210)

$$\tan\beta = -\tan\alpha\frac{\cos(i+r)}{\cos(i-r)},$$

and

$$\tan\gamma = \tan\alpha\sec(i-r).$$

Similarly both theories give the same expression for the amplitude of a ray refracted through a parallel plate, so that experiments on the intensity of the light reflected from, or refracted through, a plate will not decide between them.

**217. Minimum Reflecting Power of Transparent Bodies.**—MacCullagh[1] noticed an analogy between the results obtained for metallic reflectors and those of highly refracting transparent bodies. If we take the intensity of the incident natural light to be unity, that of the reflected, according to either theory, will be proportional to

$$\frac{\sin^2(i-r)}{\sin^2(i+r)}+\frac{\tan^2(i-r)}{\tan^2(i+r)}.$$

Equating the first derived of this expression to zero, and remembering that $\sin i=\mu\sin r$, and therefore

$$\frac{dr}{di}=\frac{\cos i}{\mu\cos r},$$

we obtain finally

$$\sin i\,(\sin^2 i-\sin^2 r)\left\{1-\frac{\cos(i+r)}{\cos^3(i-r)}\right\}=0.$$

The first and second factors vanish for $i=0$, and consequently the normal incidence corresponds to either a maximum or a minimum.

The third factor vanishes if

$$\cos^3(i-r)=\cos(i+r),$$

and this will give a value of $i$ corresponding to a minimum if the third factor is negative for very small values of $i$, for then the first derived will be negative, and the intensity of the reflected light will diminish as the incidence increases from zero to some value $i$. In order that the third factor should be negative for small values of $i$ we must have

$$\cos^3(i-r)<\cos(i+r),$$

or

$$1-\tfrac{3}{2}(i-r)^2<1-\tfrac{1}{2}(i+r)^2,$$

if $i$ and $r$ are supposed very small. Hence

$$(i+r)^2<3(i-r)^2,$$

or

$$(\mu+1)^2<3(\mu-1)^2,$$

since $i=\mu r$. Therefore

$$\mu^2-4\mu+1>0.$$

The roots of the equation $\mu^2-4\mu+1=0$ are $2+\sqrt{3}$ and $2-\sqrt{3}$, of which the latter may be discarded (being less than unity). Hence any value of $\mu$ greater than $2+\sqrt{3}$ will satisfy the above conditions, and the incidence corresponding to the minimum will be given by

$$\cos^3(i-r)=\cos(i+r).$$

---

[1] *Trans. Royal Irish Academy*, vol. xxviii. pt. i.

## CHAPTER XIV

### METALLIC REFLECTION OF POLARISED LIGHT

**218. Partial Polarisation by Reflection in General.**—Malus, who discovered the polarisation of light by reflection from glass, observed that natural light is never completely polarised by reflection from a metallic surface. The laws deduced from the theory of Fresnel are therefore not applicable to metals or highly refracting substances. M. Jamin, who investigated the question, found that only a few substances completely polarised light by reflection, that the angle of incidence at which this occurred was $\tan^{-1}(\mu)$, and for these substances $\mu = 1{\cdot}46$. For all other substances there is an angle of *maximum polarisation* determined by the equation

$$\tan i = \mu,$$

instead of an angle of *total* polarisation, and this angle is termed the *polarising angle.*

Malus was of opinion that common light is never polarised by reflection from metals, but in 1813 Brewster corrected this error, and showed that the reflected light was partially polarised, the amount of polarisation depending on the incidence and passing through a maximum at a certain angle.

Biot verified the observations of Brewster, and remarked that if light is partially polarised by one reflection at a metallic surface, it ought to be completely polarised by a sufficient number of reflections taking place at the same angle of incidence. In 1830 Brewster[1] found that when plane-polarised light is reflected from a metallic surface it remains plane-polarised if the incident ray is polarised in or perpendicular to the plane of incidence, but in any other case the light is partially "*depolarised*" by the reflection.

**219. Change of Phase and Elliptic Polarisation by Metallic Reflection.**—When the incident light is polarised in any azimuth $a$, it

[1] Brewster, *Phil. Trans.* 1830, p. 287.

may be replaced by two components, one parallel to and the other at right angles to the plane of incidence. Now these components may become altered in two respects by reflection. In the first place, their amplitudes may be changed, and secondly, their phases may be altered. The change of amplitude alone merely alters the plane of polarisation or rotates it through a certain angle, the reflected light remaining plane-polarised; hence if the reflected light is not plane-polarised, the phases of the component vibrations must have been changed by the reflection, and changed by different amounts. If this view be correct the depolarisation observed by Brewster is none other than elliptic polarisation arising from the change of phase introduced by reflection between the components parallel and perpendicular to the plane of incidence, and that this is so is supported by many experiments.

Thus when plane-polarised light suffers reflection at a metallic surface it experiences in general a change of phase by the reflection, but the change is less for light polarised at right angles to, than for light polarised in, the plane of incidence. Consequently if the incident light be polarised in or at right angles to the plane of incidence, the reflected light will remain plane-polarised in the same plane; but if it be polarised in any other plane the reflected components in and perpendicular to the plane of reflection will differ in phase, and the reflected light will in general be elliptically polarised. This difference of phase increases from zero at normal to $\pi$ at grazing incidence.

In the case of light polarised in the plane of incidence the reflected light increases in intensity from normal to grazing incidence when it is all reflected; but when the incident light is polarised at right angles to the plane of incidence the reflected ray diminishes in intensity from normal incidence to a certain angle at which it becomes a minimum, it then increases again to grazing incidence. This minimum is little marked for silver, but is very decided in the case of steel and certain metallic oxides.

Thus when natural light is reflected from a metallic surface, we may suppose it replaced by two equal components, one in and the other perpendicular to the plane of incidence. These components will be reflected in different amounts, and the reflected beam will be partially polarised. The excess of one of these components over the other is greatest at the polarising angle.

**220. M. Jamin's Experiments.**—A method of observation, susceptible of great accuracy, has been employed by M. Jamin[1] to

[1] *Annales de Chimie et de Physique*, third series, tom. xix. p. 296.

measure the intensity of the light reflected from metals. The principle of his method is the comparison of the light reflected from the metal with that reflected from glass. He employed a plane mirror, one half of which was metal and the other half glass. Light polarised by a Nicol's prism fell upon this mirror, and the reflected light was examined by a doubly refracting analyser. Two images are formed by the analyser, each of which consists of two parts, one formed by reflection from the glass, the other from the metal. The latter half is generally coloured, so that it is easily distinguished.

If the incident light be polarised in the plane of incidence, and if the principal plane of the analyser make an angle $\alpha$ with the plane of incidence, the intensities of the ordinary and extraordinary images will be measured by

| | Ordinary. | Extraordinary. |
|---|---|---|
| Glass . . . | $b^2 \cos^2 \alpha$, | $b^2 \sin^2 \alpha$, |
| Metal . . . | $m^2 \cos^2 \alpha$, | $m^2 \sin^2 \alpha$, |

where $b$, according to Fresnel's formulæ, is equal to $-a\frac{\sin(i-r)}{\sin(i+r)}$ and $m$ is a coefficient for the metal.

To determine $m$ we might seek the case when the two halves of the images are of equal intensity, and we would then have

$$m^2 = b^2 = a^2 \frac{\sin^2(i-r)}{\sin^2(i+r)}.$$

But as this incidence does not exist, and since the ordinary image of one part of the mirror and the extraordinary of the other vary from zero to a maximum in inverse senses, there will always be two values of $\alpha$, which make the ordinary image of one half equal to the extraordinary of the other, and we will have

$$m^2 \cos^2 \alpha_1 = b^2 \sin^2 \alpha_1, \quad \text{or } m^2 \sin^2 \alpha_2 = b^2 \cos^2 \alpha_2,$$

from which we obtain

$$m^2 = b^2 \tan^2 \alpha_1 = b^2 \cot^2 \alpha_2,$$

the angles $\alpha_1$ and $\alpha_2$ are therefore complementary. In practice both $\alpha_1$ and $\alpha_2$ are determined, and one observation corrects the other. The colouring in the image of the metal renders it difficult to say when the intensities of two images compared are equal, and besides this the images are not very close to each other.

For light polarised at right angles to the plane of incidence we have similarly

$$m'^2 = b'^2 \tan^2 \alpha'_1 = b'^2 \cot^2 \alpha'_2,$$

where

$$b'^2 = a^2 \frac{\tan^2(i-r)}{\tan^2(i+r)}.$$

To calculate $m$ it is necessary to know the index of refraction of the glass. This M. Jamin determined by the rotation of the plane of polarisation caused by reflection. If the azimuth of the incident light be $\alpha$, and that of the reflected $\beta$, we have

$$\tan\beta = \tan\alpha \frac{\cos(i+r)}{\cos(i-r)}.$$

Consequently if $\alpha = 45^\circ$ we have

$$\tan\beta = \frac{\cos(i+r)}{\cos(i-r)} = \frac{1 - \tan i \tan r}{1 + \tan i \tan r},$$

or

$$\tan i \tan r = \frac{1 - \tan\beta}{1 + \tan\beta} = \tan(45^\circ - \beta),$$

or

$$\tan r = \tan(45^\circ - \beta)\cot i.$$

This equation determines the angle $r$, and the refractive index is found from the equation $\sin i = \mu \sin r$.

According to Jamin's experiments the difference of phase is small from normal incidence to an angle a little less than the polarising angle; it then increases rapidly, reaches 90° at the polarising angle, and becomes very nearly equal to $\pi$ at an angle a little greater than the polarising angle. The change of phase therefore does not occur suddenly at the polarising angle, but takes place continuously and rapidly in the neighbourhood of this angle. According to Fresnel's theory we should say that the difference of phase between the two reflected components is zero up to the angle of maximum polarisation, that there it suddenly changes to $\pi$, and remains so up to grazing incidence.

A verification of the elliptic polarisation of light by metallic reflection was observed by Brewster. He found that if a ray of plane-polarised light be reflected twice at the same angle, but in perpendicular planes, from two similar metallic surfaces, the ray will be again plane-polarised after the two reflections. De Sénarmont interpreted this experiment mathematically as follows :—

Let the reflection change the component parallel to the plane of incidence by altering its amplitude in the ratio $m : 1$ and its phase by $\delta$, and let the corresponding quantities for the other component be $m'$ and $\delta'$. Then the incident and reflected vibrations in Brewster's experiment, in and perpendicular to the plane of incidence respectively, are

| Incident. | Once reflected. | Twice reflected. |
|---|---|---|
| $a \sin \omega t$, | $am \sin(\omega t + \delta)$, | $amm' \sin(\omega t + \delta + \delta')$, |
| $a' \sin \omega t$, | $a'm' \sin(\omega t + \delta')$, | $a'mm' \sin(\omega t + \delta + \delta')$. |

Hence the phases of the twice reflected components are the same, and the ratio of their amplitudes is the same as originally. The reflected beam is therefore plane-polarised in the primitive plane.

**221. Multiple Reflections.**—In the case of light reflected many times in the same plane at the same incidence, if the reflected light is *plane*-polarised the relative difference of phase introduced by reflection may be made some multiple of $\pi$. If the incident light be polarised in azimuth $\alpha$, its components parallel and perpendicular to the plane of incidence are $a \cos \alpha$ and $a \sin \alpha$, consequently if each reflection changes the amplitude of one vibration in the ratio $m : 1$ and the other in the ratio $m' : 1$, we have

| Incident light. | Reflected $n$ times. |
|---|---|
| $a \cos \alpha \sin \omega t$, | $(a \cos \alpha) m^n \sin(\omega t + n\delta)$, |
| $a \sin \alpha \sin \omega t$, | $(a \sin \alpha) m'^n \sin(\omega t + n\delta')$. |

But if the light after $n$ reflections be plane-polarised in azimuth $\beta$, we have $n\delta = n\delta' + \kappa\pi$, where $\kappa$ is a whole number, and

$$\tan \beta = \frac{b'}{b} = \left(\frac{m'}{m}\right)^n \tan \alpha.$$

Therefore

$$\frac{m'}{m} = \left(\frac{\tan \beta}{\tan \alpha}\right)^{\frac{1}{n}},$$

also if R denote the reflecting power of the surface, which may be measured photometrically for the angle of incidence used above, we have [1]

$$R^2 = \tfrac{1}{2}(m^2 + m'^2),$$

and we have thus two equations to determine $m$ and $m'$.

M. Jamin employed the method of multiple reflection to determine the change of phase produced by a single reflection at a metallic surface for any angle of incidence. Two parallel plates of metal were placed at the centre of a circle (Fig. 175), one of them fixed and the other movable by a micrometer screw so that the distance between the plates could be altered at pleasure.

The reflecting plates were initially placed at such a distance apart that the light suffered only two reflections, and the incidence was then arranged, by so turning the support on which they were placed, that

[[1] The square of the reflecting power for plane-polarised light is $m^2 \cos^2 \alpha + m'^2 \sin^2 \alpha + 2mm' \sin \alpha \cos \alpha \cos(\delta - \delta')$. For ordinary light we may neglect the last term and take the mean values of $\cos^2 \alpha$ and $\sin^2 \alpha$. Thus $R^2 = \frac{1}{2}(m^2 + m'^2)$.]

the reflected light was plane-polarised. At this incidence a difference of phase of 90° is produced at each reflection. The plates were then brought nearer (Fig. 176) so that the light was reflected 4, 6, 8 . . . times, and by turning the support gradually incidences were found for which the reflected light was plane-polarised.

The phase difference introduced at any incidence is easily calculated, for if the light suffers $n$ reflections and if $\delta$ be the change of phase at each reflection, then $n\delta$ is equal to some multiple of $\pi$ when the emergent beam is plane-polarised. Thus if with $n$ reflections the light is plane-polarised at angles of incidence $i_1$, $i_2$, $i_3$, etc., successively increasing from the normal, the change of phase for $i_1$ is $\delta_1 = \frac{\pi}{n_1}$, for $i_2$ we have $\delta_2 = \frac{2\pi}{n_2}$, while $\delta_3 = \frac{3\pi}{n_3}$, etc.

The equation $\tan i = \mu$, which determines, according to the law of Brewster, the angle of maximum polarisation, indicates that this angle is different for the different colours, increasing with $\mu$—that is, from Metallic index.

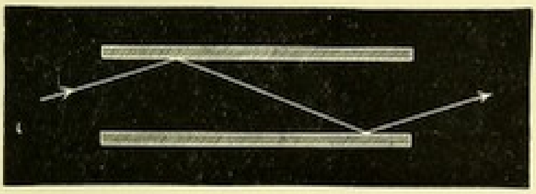


Fig. 175.

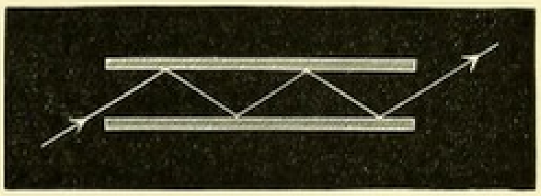


Fig. 176.

the red to the violet. M. Jamin has found, however, that the reverse is the case in metallic reflection, the angles of maximum polarisation decreasing from the red to the violet. If therefore the above equation applies to metals, we must admit that for these substances the index of refraction diminishes from the red to the violet—that is, they exhibit anomalous dispersion.

**222. MacCullagh's Theory.**—The following investigation of the intensity, and change of phase, of the light reflected from metallic surfaces may be regarded as an exercise on the formulæ derived by Fresnel for the reflection and refraction of polarised light. Depending, like Fresnel's doctrine of total reflection, on a hypothetical interpretation of an imaginary formula, it has no pretension to any physical basis, but may be interesting to any one who is curious to pursue the subject further.

In the case of metallic reflection, like that of total reflection at the surface of a transparent substance, there is no refracted wave, or at least the effect of the refracted wave is zero at a very small distance from the surface.[1]

[1] According to Fresnel in the case of total reflection a vibratory motion exists in the second medium very near the surface (*Œuvres Complètes*, i. pp. 447, 767).

In the case of total reflection we have seen that the amplitude of the reflected wave assumes the form

$$P + Q\sqrt{-1},$$

and Fresnel's interpretation of this is that a change of phase occurs, measured by $\tan\delta = Q/P$.

MacCullagh,[1] pursuing the same track, assumes that the sine of the angle of refraction in the case of metallic reflection is of the imaginary form

$$\sin r = \frac{\sin i}{m}(\cos\chi + \sqrt{-1}\sin\chi),$$

the quantities $m$ and $\chi$ being indeterminate. For transparent media $m = \mu$ and $\chi = 0$.

Since $\sin^2 r + \cos^2 r = 1$ it follows that $\cos r$ is also imaginary, and of the form

$$\cos r = \frac{\cos i}{m'}(\cos\chi' + \sqrt{-1}\sin\chi').$$

Now for light polarised in the plane of incidence the theory of Fresnel gives

$$b = -a\frac{\sin(i-r)}{\sin(i+r)} = a\frac{\dfrac{\sin r}{\sin i} - \dfrac{\cos r}{\cos i}}{\dfrac{\sin r}{\sin i} + \dfrac{\cos r}{\cos i}}.$$

Substituting for $\sin r/\sin i$ and $\cos r/\cos i$ from the foregoing expressions the formula for $b$ becomes

$$b = a\frac{(m'\cos\chi - m\cos\chi') + (m'\sin\chi - m\sin\chi')\sqrt{-1}}{(m'\cos\chi + m\cos\chi') + (m'\sin\chi + m\sin\chi')\sqrt{-1}},$$

or multiplying by the conjugate form of the denominator we find

$$b = a\frac{m'^2 - m^2 + 2mm'\sin(\chi - \chi')\sqrt{-1}}{m^2 + m'^2 + 2mm'\cos(\chi - \chi')}.$$

This expression is of the form $P + Q\sqrt{-1}$, and interpreting it as in the case of total reflection, we find for the intensity of the reflected light $P^2 + Q^2$, or

$$I = a^2\frac{(m'^2 - m^2)^2 + 4m^2m'^2\sin^2(\chi - \chi')}{[m^2 + m'^2 + 2mm'\cos(\chi - \chi')]^2},$$

$$= a^2\frac{m^2 + m'^2 - 2mm'\cos(\chi - \chi')}{m^2 + m'^2 + 2mm'\cos(\chi - \chi')},$$

[1] MacCullagh, *Proc. Roy. Irish Acad.* vol. i. p. 2, 1836.

and for the change of phase

$$\tan\delta = \frac{2mm'\sin(\chi - \chi')}{m'^2 - m^2}.$$

To determine $m'$ and $\chi'$ in terms of $m$ and $\chi$, we have

$$\sin^2 r + \cos^2 r = 1,$$

or

$$\frac{\sin^2 i}{m^2}\cos 2\chi + \frac{\cos^2 i}{m'^2}\cos 2\chi' + \left(\frac{\sin^2 i}{m^2}\sin 2\chi + \frac{\cos^2 i}{m'^2}\sin 2\chi'\right)\sqrt{-1} = 1.$$

Therefore

$$\frac{\sin^2 i}{m^2}\cos 2\chi + \frac{\cos^2 i}{m'^2}\cos 2\chi' = 1,$$

and

$$\frac{\sin^2 i}{m^2}\sin 2\chi + \frac{\cos^2 i}{m'^2}\sin 2\chi' = 0.$$

Hence

$$\sin 2\chi' = -\frac{m'^2}{m^2}\tan^2 i \sin 2\chi$$

and

$$\cos 2\chi' = +\frac{m'^2(m^2 - \sin^2 i\cos 2\chi)}{m^2\cos^2 i},$$

or squaring and adding the expressions for $\sin 2\chi'$ and $\cos 2\chi'$, we have

$$1 = \frac{m'^4}{m^4\cos^4 i}(m^4 + \sin^4 i - 2m^2\sin^2 i\cos 2\chi),$$

or

$$m'^2 = \frac{m^2\cos^2 i}{\sqrt{m^4 + \sin^4 i - 2m^2\sin^2 i\cos 2\chi}} = \frac{m^2\cos^2 i}{D^2},$$

where

$$D^2 = \sqrt{m^4 + \sin^4 i - 2m^2\sin^2 i\cos 2\chi}.$$

Again

$$\tan 2(\chi - \chi') = \frac{\tan 2\chi - \tan 2\chi'}{1 + \tan 2\chi\tan 2\chi'} = \frac{\tan 2\chi + \dfrac{\sin^2 i\sin 2\chi}{m^2 - \sin^2 i\cos 2\chi}}{1 - \tan 2\chi\dfrac{\sin^2 i\sin 2\chi}{m^2 - \sin^2 i\cos 2\chi}}$$

$$= \frac{m^2\sin 2\chi}{m^2\cos 2\chi - \sin^2 i},$$

and

$$\sin 2(\chi - \chi') = \frac{m^2\sin 2\chi}{\sqrt{m^4 + \sin^4 i - 2m^2\sin^2 i\cos 2\chi}} = \frac{m^2\sin 2\chi}{D^2}.$$

Hence

$$I = a^2\frac{D^2 + \cos^2 i - 2D\cos i\cos(\chi - \chi')}{D^2 + \cos^2 i + 2D\cos i\cos(\chi - \chi')},$$

and

$$\tan\delta = \frac{2D\cos i\sin(\chi - \chi')}{\cos^2 i - D^2}.$$

In the case of light polarised at right angles to the plane of incidence we have

$$-b' = a\frac{\tan(i-r)}{\tan(i+r)} = a\frac{1-\dfrac{\sin r\cos r}{\sin i\cos i}}{1+\dfrac{\sin r\cos r}{\sin i\cos i}},$$

$$= a\frac{mm' - \cos(\chi+\chi') - \sqrt{-1}\sin(\chi+\chi')}{mm' + \cos(\chi+\chi') + \sqrt{-1}\sin(\chi+\chi')},$$

$$= a\frac{m^2m'^2 - 1 - 2mm'\sin(\chi+\chi')\sqrt{-1}}{m^2m'^2 + 1 + 2mm'\cos(\chi+\chi')},$$

which is of the form $P + Q\sqrt{-1}$. Hence

$$I' = a^2\frac{m^2m'^2 + 1 - 2mm'\cos(\chi+\chi')}{m^2m'^2 + 1 + 2mm'\cos(\chi+\chi')},$$

and for the change of phase

$$\tan\delta' = -\frac{2mm'\sin(\chi+\chi')}{m^2m'^2 - 1}.$$

Or substituting for $m'$, we finally obtain

$$I' = a^2\frac{m^4\cos^2 i + D^2 - 2Dm^2\cos i\cos(\chi+\chi')}{m^4\cos^2 i + D^2 + 2Dm^2\cos i\cos(\chi+\chi')},$$

and

$$\tan\delta' = -\frac{2Dm^2\cos i\sin(\chi+\chi')}{m^4\cos^2 i - D^2}.$$

**223. Simplification in the case of Metals.**—Since at normal incidence the light reflected from a polished metallic surface is very nearly equal to the incident light, it follows that the quantity

$$\frac{m^2 + 1 - 2m\cos\chi}{m^2 + 1 + 2m\cos\chi}$$

should be very nearly equal to unity. For when $i = 0$, $D^2 = m^2$, and $\sin 2(\chi - \chi') = \sin 2\chi$ or $\chi' = 0$.

Hence $m$ must be very great, and if we neglect $\chi'$, we have

$$m' = \cos i/\cos r.$$

Substituting this value of $m'$ in the expressions for I and $\delta$, we have

$$I = \frac{m^2 + \dfrac{\cos^2 i}{\cos^2 r} - 2m\dfrac{\cos i}{\cos r}\cos\chi}{m^2 + \dfrac{\cos^2 i}{\cos^2 r} + 2m\dfrac{\cos i}{\cos r}\cos\chi}, \qquad \tan\delta = \frac{2m\dfrac{\cos i}{\cos r}\sin\chi}{\dfrac{\cos^2 i}{\cos^2 r} - m^2},$$

and

$$I' = \frac{1 + m^2\dfrac{\cos^2 i}{\cos^2 r} - 2m\dfrac{\cos i}{\cos r}\cos\chi}{1 + m^2\dfrac{\cos^2 i}{\cos^2 r} + 2m\dfrac{\cos i}{\cos r}\cos\chi}, \qquad \tan\delta' = -\frac{2m\dfrac{\cos i}{\cos r}\sin\chi}{m^2\dfrac{\cos^2 i}{\cos^2 r} - 1}.$$

# CHAPTER XV

## INTERFERENCE OF POLARISED LIGHT—COLOURS OF THIN CRYSTALLINE PLATES

### 1. *Parallel Plane-Polarised Light*

**224. Introductory Statement.**—We now proceed to the study of the phenomena which occur when polarised light is transmitted through thin plates of doubly refracting substances. The first discoveries in this region were made by Arago[1] in 1811. Placing by chance a thin plate of mica in the path of a pencil of plane-polarised light (the blue light of the sky) and examining it through a doubly refracting prism (Iceland spar), he observed that both the ordinary and the extraordinary images were richly coloured. In general, when plane-polarised light is transmitted through a thin plate of any doubly refracting substance and then examined by means of a doubly refracting analyser, both images are richly coloured, and if they overlap their common portion appears white, which shows that the colours of the images are complementary.

If plane-polarised light be received by a Nicol's prism, or other analyser, we know that in one position of the Nicol the light is refused transmission. The Nicol being set in this position, if a thin plate of a crystal be introduced across the path of the light, the capability of transmission through the Nicol is suddenly restored, and a portion of the light is transmitted which depends on the position of the interposed crystal. For this reason the light was said to be "depolarised" by the crystal, and by means of this property the doubly refracting structure of many substances was detected by Malus, where the separation of the ray was too small to be observed directly.

A medium originally isotropic may acquire the power of double refraction when subjected to strain, and if the strain be homogeneous the optical properties of the substance are similar to those of a natural crystal, the principal axes of the wave surface coinciding with those of

[1] Arago, *Œuvres Complètes*, tom. x. p. 36.

the strain. A feeble doubly refracting power is conferred on glass by bending or straining it. Unequal heating produces the same effect, as it is accompanied by expansion, which gives rise to internal strain. This double refraction may be detected by examining the glass between crossed Nicols, and so delicate is this test that it is difficult to find large pieces of glass so free from internal strain as to show no revival of light when so examined.

**225. Intensity of Illumination at any Point.** — Let us now see how far the physical theory accounts for these appearances. In the first place, there are three essential conditions for their production.

(1) The polarisation of the incident light.

(2) The interposition of a thin crystalline plate.

(3) The action of an analyser on the light after passing through the plate.

Let the principal plane of the polarising Nicol be parallel to OP (Fig. 177) and let the principal plane of the analysing Nicol be parallel to OA. Then since it is the extraordinary ray that is transmitted through the Nicol, and since by Fresnel's hypothesis the vibrations of this ray are in the principal plane, it follows that the incident vibration at O will be parallel to OP. On entering the plate it becomes broken up into two others, polarised at right angles to each other, one parallel to OX and the other parallel to OY, where OX and OY are two determinate rectangular directions in the crystal. Hence if the incident vibration be $y = a \sin \omega t$ it gives rise to

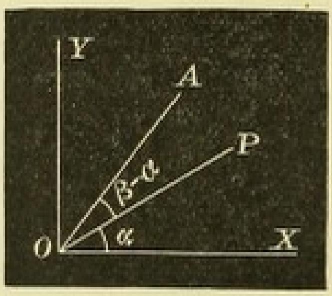



Fig. 177.

$$a \cos \alpha \sin \omega t, \quad \text{and } a \sin \alpha \sin \omega t$$

along OX and OY where $POX = \alpha$.

As these waves travel through the plate with different velocities they will be unequally retarded, and consequently on emergence they will differ in phase by an amount $\delta$. The transmitted vibrations therefore take the form

$$a \cos \alpha \sin \omega t, \quad \text{and } a \sin \alpha \sin (\omega t + \delta).$$

On reaching the analyser these vibrations become resolved parallel to its principal plane OA. If therefore $AOX = \beta$ we have two vibrations parallel to the principal plane of the analyser: the component $a \cos \alpha \sin \omega t$ along OX gives $a \cos \alpha \cos \beta \sin \omega t$ along OA, and the component $a \sin \alpha \sin (\omega t + \delta)$ gives $a \sin \alpha \sin \beta \sin (\omega t + \delta)$ along OA, and these compound into a resultant vibration

$$y = a \cos \alpha \cos \beta \sin \omega t + a \sin \alpha \sin \beta \sin (\omega t + \delta),$$

parallel to OA. The intensity of the resultant vibration is therefore given by (Art. 43, chap. ii.) the equation

$$\begin{aligned} \mathrm{I} &= a^2\{\cos^2 \alpha \cos^2 \beta + \sin^2 \alpha \sin^2 \beta + 2 \sin \alpha \cos \alpha \sin \beta \cos \beta \cos \delta\}, \\ &= a^2\{\cos^2 \alpha \cos^2 \beta + \sin^2 \alpha \sin^2 \beta + 2 \sin \alpha \cos \alpha \sin \beta \cos \beta (1 - 2 \sin^2 \tfrac{1}{2}\delta)\}, \\ &= a^2\{(\cos \alpha \cos \beta + \sin \alpha \sin \beta)^2 - 4 \sin \alpha \cos \alpha \sin \beta \cos \beta \sin^2 \tfrac{1}{2}\delta\}. \end{aligned}$$

Hence finally

$$\mathrm{I} = a^2\{\cos^2 (\alpha - \beta) - \sin 2\alpha \sin 2\beta \sin^2 \tfrac{1}{2}\delta\},$$

where $\alpha - \beta$ is the angle between the principal planes of the polariser and analyser.

If the analyser be merely a doubly refracting rhomb two images will be presented. The above expression refers to the extraordinary image, and it is seen in the same manner that the intensity of the ordinary image is

$$\mathrm{I}_0 = a^2\{\sin^2 (\alpha - \beta) + \sin 2\alpha \sin 2\beta \sin^2 \tfrac{1}{2}\delta\},$$

which shows that it is complementary to the extraordinary. We shall confine our attention to the extraordinary, as it is that furnished by a single image analyser such as Nicol's prism.

In the case of white light $\delta$ will be different for the various wave lengths,[1] and if $a$ also varies with the wave length, the general expression for the intensity will be

$$\mathrm{I} = \cos^2 (\alpha - \beta)\, \Sigma a^2 - \sin 2\alpha \sin 2\beta\, \Sigma a^2 \sin^2 \tfrac{1}{2}\delta.$$

The first term being independent of $\delta$ will have no effect in producing colour in the image, but in the second term $\delta$ will depend on the wave length, and consequently the different colours will enter it in different amounts. If the incident light be white the transmitted light will in general consist of two parts, one of which is white, depending on the first term, and the other more or less coloured, arising from the second. With a given plate the combined rotation of the Nicols, or the rotation of the plate round its normal, will affect all the colours in the same proportion, and consequently the tint of the second term will remain the same, but its intensity will vary as $\sin 2\alpha \sin 2\beta$ varies. Now

The colc term.

[1] In some crystals the dispersion sensibly modifies the relative retardation as dependent on the wave length. Herschel (Art. "Light," *Ency. Metropolitana*, § 915) observed that the rings exhibited by a common variety of uniaxal apophyllite were approximately achromatic, indicating that $\delta$ was almost independent of $\lambda$, and under these circumstances a very great number of rings may become visible.

$\sin 2\alpha \sin 2\beta$ may be either positive or negative, and for this reason the resultant colour of the plate may be either of two different tints. For example, when $\sin 2\alpha \sin 2\beta$ is positive (which will hold as the plate is rotated from $\alpha = 0°$ to $\beta = 90°$), the resultant light will consist of a certain quantity of white light, from which a varying amount of light of a given colour is *subtracted*, and when $\sin 2\alpha \sin 2\beta$ is negative (which will hold from $\beta = 90°$ to $\alpha = 90°$), the resultant will consist of a given quantity of white light, to which a varying quantity of light of a given colour is added. In each case the resultant tint remains unaltered as the plate is rotated (until $\sin 2\alpha \sin 2\beta$ changes sign), except in so far as it becomes more or less diluted by the greater or less admixture of white light arising from the first term of I.

The colouring, depending on the second term, will be most marked when $\alpha - \beta = 90°$, and least marked when $\alpha - \beta = 0$, the corresponding values of I being

$$\mathrm{I} = \Sigma a^2 \sin^2 2\alpha \sin^2 \tfrac{1}{2}\delta \quad \text{(colour most marked)},$$
$$\mathrm{I} = \Sigma a^2 (1 - \sin^2 2\alpha \sin^2 \tfrac{1}{2}\delta) \quad \text{(colour least marked)}.$$

In the former case the field of the analyser would be dark if the plate were removed. In both cases the appearances are most marked when $\sin 2\alpha = 1$, or $\alpha = 45°$—that is, when the principal planes of the polariser and analyser bisect the angles between the principal planes of the plate.

When the light falling on the thin plate is not polarised there is no exhibition of colour in the field of the analyser, and the images are white if the incident light be ordinary white light. The light suffers double refraction in the crystalline plate, but on account of its small thickness there is no visible separation of the rays, they consequently combine and still preserve the property of giving two images of equal intensity in a doubly refracting analyser, the intensity of each image being half that of the incident light.

Complementary relation.

*Cor.* 1.—The intensity in any position $\beta$ of the analyser is complementary to that in the perpendicular position $\beta + 90°$. For if $\mathrm{I}_1$ and $\mathrm{I}_2$ denote these two intensities, we have

$$\mathrm{I}_1 = a^2\{\cos^2(\alpha - \beta) - \sin 2\alpha \sin 2\beta \sin^2 \tfrac{1}{2}\delta\},$$

and changing $\beta$ into $90° + \beta$, we have

$$\mathrm{I}_2 = a^2\{\sin^2(\alpha - \beta) + \sin 2\alpha \sin 2\beta \sin^2 \tfrac{1}{2}\delta\},$$

therefore

$$\mathrm{I}_1 + \mathrm{I}_2 = a^2,$$

or the sum of the two intensities is equal to that of the incident light.

The effect of rotating the analyser through 90° is therefore to change the intensity, and also the tint, into its complementary. If the analyser be merely a doubly refracting rhomb, the two images will exhibit these complementary tints, and if the analysing rhomb be removed altogether these complementary images will become superposed, and we have a single image without colour.

*Cor.* 2.—For a given relative position of the Nicols—that is, for $\alpha - \beta = \text{constant} = \gamma$—the intensity of the coloured component will vary as the plate is rotated round the normal to its plane, and the image will be uncoloured when Achromatic conditio

$$\sin 2\alpha \sin 2\beta = 0,$$

that is, when $\alpha = 0°$ or $90°$, and when $\beta = 0°$ or $90°$. There are consequently four positions in which the image is achromatic, viz. when the principal section of the plate is parallel or perpendicular to the principal plane of the polariser or analyser, and the intensity of the achromatic image is

$$I = a^2 \cos^2 \gamma.$$

This will be a maximum when $\gamma = 0°$—that is, when $\alpha = \beta$, or when the polariser and analyser are parallel, and zero when $\gamma = 90°$,—that is, when the Nicols are crossed. In both these cases the four positions giving an achromatic image obviously reduce to two.

*Cor.* 3.—If $\alpha = \beta$, or the principal planes of the Nicols are parallel, we have for the intensity of the transmitted light

$$I = a^2(1 - \sin^2 2\alpha \sin^2 \tfrac{1}{2}\delta).$$

*Cor.* 4.—If in addition $\alpha = 0°$ or $90°$, we have

$$I = a^2,$$

or the transmitted light is equal to the incident. The plate has here no effect, as is easily understood, for if $\alpha$ is equal to 0° or 90°, the direction of vibration of the incident light is parallel to one of the possible directions in the crystal. It therefore passes through unaltered to the analyser.

If $\alpha = \beta = 45°$, $I = a^2 \cos^2 \frac{1}{2}\delta$. Hence if $\delta = (2n+1)\pi$ and $\alpha = \beta = 45°$, the transmitted light is zero.

*Cor.* 5.—For any given positions of the Nicols and plate the intensity is a maximum when $\sin \frac{1}{2}\delta = 0$ and a minimum when $\sin \frac{1}{2}\delta = \pm 1$ if $\sin 2\alpha \sin 2\beta$ is positive, or

$$I = a^2 \cos^2(\alpha - \beta), \quad \delta = 2n\pi \qquad \text{(max.)},$$
$$I = a^2\{\cos^2(\alpha - \beta) - \sin 2\alpha \sin 2\beta\}, \quad \delta = (2n+1)\pi \qquad \text{(min.)}.$$

But if $\sin 2\alpha \sin 2\beta$ is negative, the case is reversed. The difference of phase $\delta$ is determined by the thickness of the plate traversed and by the wave length of the light under investigation.

*Cor.* 6.—If $\alpha - \beta = 90^\circ$, so that the Nicols are crossed, we have $\sin 2\beta = -\sin 2\alpha$, and

$$I = a^2 \sin^2 2\alpha \sin^2 \tfrac{1}{2}\delta.$$

So that for a given value of $\alpha$ the intensity I will be a maximum when $\delta = (2n+1)\pi$ and zero when $\delta = 2n\pi$. In the former case the intensity will be the greatest possible when $\alpha = 45^\circ$, for then $I = a^2$, and the intensity will be zero if $\alpha = 0$ or $90^\circ$.

*Cor.* 7.—To determine the phase of the resultant vibration

$$\begin{aligned} y &= a\{\cos\alpha\cos\beta\sin\omega t + \sin\alpha\sin\beta\sin(\omega t+\delta)\}, \\ &= A\sin(\omega t+\rho), \end{aligned}$$

we have

$$\begin{aligned} A\sin\rho &= a\sin\alpha\sin\beta\sin\delta, \\ A\cos\rho &= a\{\cos\alpha\cos\beta + \sin\alpha\sin\beta\cos\delta\}, \\ \therefore\ \tan\rho &= \frac{\sin\alpha\sin\beta\sin\delta}{\cos\alpha\cos\beta+\sin\alpha\sin\beta\cos\delta}. \end{aligned}$$

Hence if $\alpha$ and $\beta$ are the semi-sides, and $\delta$ the included angle, of a spherical triangle, $\rho$ is half the spherical excess (*Spher. Trig.* Art. 105).

*Cor.* 8.—The light emerging from the thin plate is in general elliptically polarised, the equation of the vibration being (Art. 47)

$$\frac{x^2}{\cos^2\alpha} + \frac{y^2}{\sin^2\alpha} - \frac{2xy\cos\delta}{\sin\alpha\cos\alpha} = a^2\sin^2\delta.$$

**226. Conditions for Interference — Transverse Vibrations.**— The phenomenon of double refraction shows that in the crystal the light is divided into two waves travelling with different velocities. On emerging therefore from a thin plate one will be retarded on the other, and in general a difference of phase will exist. Hence if light undulations were purely longitudinal, like those of sound, the emerging waves should interfere, and the thin plate alone should be sufficient to produce all the phenomena of interference and colour without either polariser or analyser. Such, however, is not the case, and it was found by Fresnel and Arago, who investigated the subject of the interference of polarised light experimentally, that two plane-polarised rays interfere and produce fringes as ordinary light only when they are polarised in the same plane and originally belonged to the same plane-polarised pencil (Art. 176).

Two rays polarised in different planes in general combine into a resultant elliptic vibration, and this is an immediate consequence of the theory of transverse vibrations. The office of the analyser in these experiments is to reduce the two waves from the crystal to the same plane of polarisation. They can then interfere with the production of colours. All the phenomena are therefore in confirmation of the theory which supposes that the vibrations of the ether which constitute light involve a component which is transverse to the direction of propagation.

**227. Calculation of the Difference of Phase—Quarter-wave Plate.**—So far we have only considered the nature of the vibration which reaches the eye from a single point O of the thin plate, or from a uniform plate transmitting parallel light. If the incident light be not parallel, and if the thickness of the plate varies from point to point, the retardation $\delta$ of one component on the other will also vary from point to point of the plate, so that the calculation of the appearance of the plate requires the determination of $\delta$ at each point.

A ray incident at an angle $i$ gives rise to two refracted rays, and the relative retardation introduced by a plate of thickness $e$ we have already found (Ex., p. 99) to be

$$\delta = e \sin i (\cot r_2 - \cot r_1) = e(\mu_2 \cos r_2 - \mu_1 \cos r_1),$$

and on the calculation of the quantity $\delta$ the determination of the character of the pattern presented in the field will depend.

For normal incidence the retardation becomes $e(\mu_2 - \mu_1)$, and if the thickness of the plate is so adjusted that

$$e(\mu_2 - \mu_1) = \tfrac{1}{4}\lambda,$$

the plate will be a *quarter-wave plate* for the wave length $\lambda$. Such plates are of importance in the study of circularly polarised light.

**228. Thick Plates—Colours produced by Superposition.**—The phenomenon of colour we have seen to be due to the difference of phase introduced by the thin crystalline plate, and if the plate be not thin this difference may be a great number of wave lengths, so that, as in the case of Newton's rings, the colours of different orders may come to be superposed, and the resultant light will be white.

Thus for a given position of the Nicols and plate ($\alpha$ and $\beta$ given) the intensity will be a maximum or a minimum according as $\sin^2 \frac{1}{2}\delta = 0$ or 1, assuming $\sin 2\alpha \sin 2\beta$ positive. The maximum or minimum intensity for any colour will consequently correspond to retardations $n\lambda$ and $(2n+1)\frac{1}{2}\lambda$ respectively. But if the plate is thick $n$ will be very great, so that if $\lambda$ and $\lambda'$ are two near wave lengths, we may have Over-lapping.

$$2n \cdot \tfrac{1}{2}\lambda = (2n+1)\tfrac{1}{2}\lambda',$$

and thus if the plate is dark for $\lambda'$ it will be bright for an adjacent wave length $\lambda$. It will therefore be bright for many wave lengths along the whole range of the spectrum, and will consequently appear white. When examined under a spectroscope it should, nevertheless, exhibit a spectrum crossed by many dark bands.

The tints may, however, be produced in thick plates by superposing two of them in such a manner that the ray which has the greater velocity in the first shall have the lesser velocity in the second. Thus if the crystals be uniaxal and both positive (or both negative) their principal sections should be placed at right angles, whereas these sections should be placed parallel if the crystals are of opposite denominations.

**229. Superposition of two Crystalline Plates.**—Let us now examine the appearances presented when plane-polarised light passes successively through two superposed thin crystalline plates, before being received by the analyser.

Using the notation of Art. 225, we shall take the direction of the vibration incident on the first plate to be parallel to OP, and this according to Fresnel's theory is parallel to the principal plane of the polariser when the light is polarised by transmission through a Nicol's prism. Then if OX and OY be the two directions of vibration in the first plate, the components of the light emerging from it may be written in the form

$$a \cos \alpha \sin (\omega t + \delta_1), \quad \text{and } a \sin \alpha \sin (\omega t + \delta'_1),$$

where $\delta_1$ and $\delta'_1$ are the phase retardations introduced by the first plate. When these vibrations reach the second plate they are each split up into two components, except in the particular cases in which the principal planes of the plates are parallel to each other or crossed.

For the sake of simplicity we shall first consider the case in which the principal planes of the plates are parallel to each other. In this case the foregoing vibrations traverse the second plate without subdivision, and on emerging from it they may be written in the form

Principal planes parallel.

$$a \cos \alpha \sin (\omega t + \delta_1 + \delta_2), \quad \text{and } a \sin \alpha \sin (\omega t + \delta'_1 + \delta'_2),$$

where $\delta_2$ and $\delta'_2$ are the phase retardations introduced by the second plate and correspond to $\delta_1$ and $\delta'_1$ respectively in the first. The whole retardation of one ray is $\delta_1 + \delta_2$, while that of the other is $\delta'_1 + \delta'_2$; hence falling upon the analyser we have two rectangular vibrations differing in phase by an amount $(\delta_1 + \delta_2) - (\delta'_1 + \delta'_2)$, and the intensity of the light vibrating in the direction of the principal plane of the analyser is consequently found from the expression of Art. 225 by merely replacing the quantity $\delta$ by the quantity $(\delta_1 + \delta_2) - (\delta'_1 + \delta'_2)$. This is the extraordinary image, and therefore its intensity is given by the expression

$$I_e = a^2\{\cos^2(\alpha-\beta) - \sin 2\alpha \sin 2\beta \sin^2 \tfrac{1}{2}(\delta_1+\delta_2-\delta'_1-\delta'_2)\}.$$

Similarly the intensity of the ordinary image, if transmitted by the analyser, will be

$$I_o = a^2\{\sin^2(\alpha-\beta) + \sin 2\alpha \sin 2\beta \sin^2 \tfrac{1}{2}(\delta_1+\delta_2-\delta'_1-\delta'_2)\}.$$

Thus when the principal planes of the plates are parallel, the combination acts as a single plate of thickness equivalent to the joint thicknesses of the two. When the plates are crossed the phase retardation of one ray is obviously $\delta_1 + \delta'_2$, while that of the other is $\delta'_1 + \delta_2$; consequently the difference of phase on reaching the analyser is $(\delta_1 + \delta'_2) - (\delta'_1 + \delta_2)$, and the intensity of the extraordinary image is

Plates crossed.

$$I_e = a^2\{\cos^2(\alpha-\beta) - \sin 2\alpha \sin 2\beta \sin^2 \tfrac{1}{2}(\delta_1+\delta'_2-\delta'_1-\delta_2)\},$$

while the intensity of the ordinary image is given by the complementary expression

$$I_o = a^2\{\sin^2(\alpha-\beta) + \sin 2\alpha \sin 2\beta \sin^2 \tfrac{1}{2}(\delta_1+\delta'_2-\delta'_1-\delta_2)\}.$$

In the same way the expression for the intensity may be written down at once when any number of plates are superposed with their principal planes either parallel or crossed, or when some of them are parallel while others are crossed, for we have merely to replace the quantity $\delta$ in the formulæ of Art. 225 by a corresponding quantity $\Sigma\delta - \Sigma\delta'$, where $\Sigma\delta$ is the whole phase retardation of one of the rays emerging from the system of plates, and $\Sigma\delta'$ the whole phase retardation of the other.

In the more general case in which the principal planes of the plates are inclined to each other at an angle, the calculation is more tedious but presents no serious difficulty. Thus, using the same notation, if the vibrations emerging from the first plate parallel to OX and OY are

Principal planes inclined.

$$a \cos \alpha_1 \sin(\omega t + \delta_1), \quad \text{and } a \sin \alpha_1 \sin(\omega t + \delta'_1);$$

then if the directions of vibration OX′ and OY′ in the second crystal be inclined at an angle $\alpha_2$ to those in the first, viz. OX and OY, the foregoing vibrations give for the vibration parallel to OX′

$$X = a \cos\alpha_1 \cos\alpha_2 \sin(\omega t + \delta_1 + \delta_2) + a \sin\alpha_1 \sin\alpha_2 \sin(\omega t + \delta'_1 + \delta_2),$$

while parallel to OY′ we have

$$Y = a \sin\alpha_1 \cos\alpha_2 \sin(\omega t + \delta'_1 + \delta'_2) - a \cos\alpha_1 \sin\alpha_2 \sin(\omega t + \delta_1 + \delta'_2).$$

Hence if the principal plane OP of the analyser makes an angle $\alpha_3$ with OX′, the component vibration parallel to OP is

$$y = X \cos \alpha_3 + Y \sin \alpha_3.$$

Substituting for X and Y and collecting the coefficients of cos $\omega t$ and sin $\omega t$ we find

$$y = P \cos \omega t + Q \sin \omega t,$$

where

$$P/a = \cos \alpha_1 \cos \alpha_2 \cos \alpha_3 \sin(\delta_1 + \delta_2) + \sin \alpha_1 \sin \alpha_2 \cos \alpha_3 \sin(\delta'_1 + \delta_2)$$
$$+ \sin \alpha_1 \cos \alpha_2 \sin \alpha_3 \sin(\delta'_1 + \delta'_2) - \cos \alpha_1 \sin \alpha_2 \sin \alpha_3 \sin(\delta_1 + \delta'_2),$$

and

$$Q/a = \cos \alpha_1 \cos \alpha_2 \cos \alpha_3 \cos(\delta_1 + \delta_2) + \sin \alpha_1 \sin \alpha_2 \cos \alpha_3 \cos(\delta'_1 + \delta_2)$$
$$+ \sin \alpha_1 \cos \alpha_2 \sin \alpha_3 \cos(\delta'_1 + \delta'_2) - \cos \alpha_1 \sin \alpha_2 \sin \alpha_3 \sin(\delta_1 + \delta'_2).$$

Consequently for the intensity of the light vibrating parallel to the principal plane of the analyser—that is, for the extraordinary image—we have

$$I_e = P^2 + Q^2.$$

Hence by taking the sum of the squares P and Q, and noticing that the coefficients of the quantities sin $(\delta_1 + \delta_2)$, etc., in the expressions for P and Q are exactly the terms which occur in the expansion of cos $(\alpha_1 - \alpha_2 - \alpha_3)$, we obtain

$$I_e = a^2\{\cos^2(\alpha_1 - \alpha_2 - \alpha_3) - \sin 2\alpha_1 \sin 2\alpha_2 \cos 2\alpha_3 \sin^2 \tfrac{1}{2}(\delta_1 - \delta'_1)$$
$$- \sin 2\alpha_1 \cos^2 \alpha_2 \sin 2\alpha_3 \sin^2 \tfrac{1}{2}(\delta_1 + \delta_2 - \delta'_1 - \delta'_2)$$
$$+ \cos 2\alpha_1 \sin 2\alpha_2 \sin 2\alpha_3 \sin^2 \tfrac{1}{2}(\delta_2 - \delta'_2)$$
$$+ \sin 2\alpha_1 \sin^2 \alpha_2 \sin 2\alpha_3 \sin^2 \tfrac{1}{2}(\delta_1 + \delta'_2 - \delta'_1 - \delta_2)\}.$$

In the same way for the intensity of the ordinary image

$$I_o = a^2\{\sin^2(\alpha_1 - \alpha_2 - \alpha_3) + \sin 2\alpha_1 \sin 2\alpha_2 \cos 2\alpha_3 \sin^2 \tfrac{1}{2}(\delta_1 - \delta'_1)$$
$$+ \sin 2\alpha_1 \cos^2 \alpha_2 \sin 2\alpha_3 \sin^2 \tfrac{1}{2}(\delta_1 + \delta_2 - \delta'_1 - \delta'_2)$$
$$- \cos 2\alpha_1 \sin 2\alpha_2 \sin 2\alpha_3 \sin^2 \tfrac{1}{2}(\delta_2 - \delta'_2)$$
$$- \sin 2\alpha_1 \sin^2 \alpha_2 \sin 2\alpha_3 \sin^2 \tfrac{1}{2}(\delta_1 + \delta'_2 - \delta'_1 - \delta_2)\}.$$

By adding these expressions together we find that the two images are complementary, for we have

$$I_o + I_e = a^2.$$

The tint of the image in either case will vary with $\alpha_1$, $\alpha_2$, and $\alpha_3$,—that is, when the analyser or either of the plates is rotated.

*Cor.*—Making $\alpha_2 = 0$, we obtain the expression for two plates with their principal sections parallel, and making $\alpha_2 = 90°$, we obtain the formulæ for two plates crossed.

The foregoing enables us, in a simple manner, to determine whether a crystal is positive or negative. For this purpose take a thin plate of the crystal and observe the tints it produces in polarised light. Now superpose on it a plate of quartz or some other crystal of known sign so that the principal sections of the two plates may be parallel. The two crystals will be of the same or contrary signs according as the new tints presented in the analyser are higher or lower in Newton's scale of colours than those afforded before the interposition of the quartz plate. Sign test.

If the principal sections of the two superposed plates are parallel or perpendicular the images presented by the analyser are not altered by interchanging the positions of the plates, but this is not generally the case when the principal sections make any other angle with each other, for the above formulæ show that if $\alpha_1$ and $\alpha_2$ be interchanged the values of $I_o$ and $I_e$ also change except when $\alpha_2 = 0^\circ$ or $90^\circ$.

**230. Projection on a Screen.**—The phenomena indicated by our theory may be observed by simply looking at the sky through two

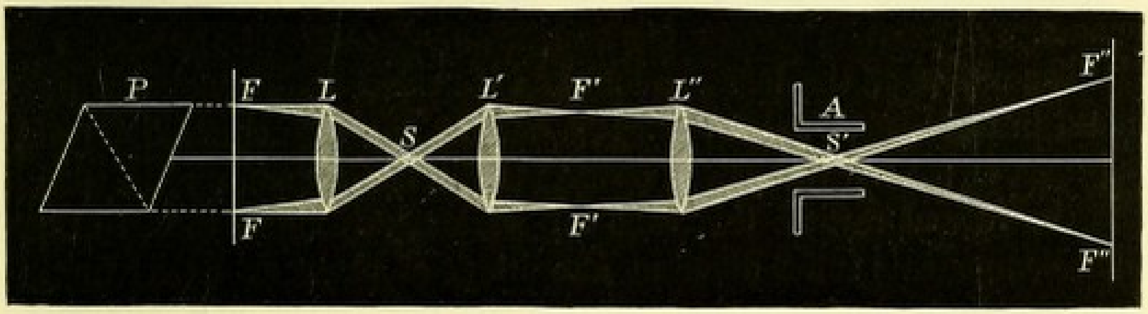



Fig. 178.

Nicol's prisms separated by a thin crystalline plate. But by projecting the images on a screen as follows, they may be observed on a larger scale and exhibited to an audience.

Let the light of the sun, reflected from a heliostat if necessary, be transmitted through a polarising Nicol or Foucault prism P (Fig. 178).

The polarised light from P falls upon a system of two highly converging lenses L and L′, having a common focus at S. It is clear that L produces an image of the sun at S and the light leaves L′ in the same condition as it enters L. If another converging lens L″ be placed in the path of the light, it will again be brought to a focus at S′, where we shall have a second image of the sun from which the light will diverge. It is at S′, where the pencil is very narrow, that the analysing prism is placed.

Now if a diaphragm be placed at FF we may regard each point of it as the origin of the small conical pencil of 16′ aperture which it receives from the sun. Any one of these pencils diverging from F will be refracted at L and travel along LS in a parallel beam and then come to a real focus F′ from which it diverges, and falling upon L″ is brought to a focus F‴ upon a screen F″F″.

If we wish to observe the tints produced by a crystalline plate when the incident light is parallel, it is placed at FF or at F′F′, the pair of lenses L and L′ in this case having obviously no effect. The incident light being polarised in an azimuth $\alpha$ to the principal section of the plate will be divided by it into two parts polarised at right angles and differing in phase. Both these parts are concentrated by L″ at S′, where the analyser is placed, and are again reduced by it to a definite polarised pencil, which paints an image on the screen F″F″. If the analyser is merely a doubly refracting prism we shall have two images on the screen which are complementary in colour, as is verified by the fact that if they are partially superposed the overlapping portion is always white, no matter how the plate is changed. By inclining the crystalline plate to the direction of the light, by turning it round a line in its plane, we alter the thickness of the plate traversed and change the difference of phase, and therefore vary the tints of the images.

The object of the lenses L and L′ is to study the phenomena produced in convergent light. For this purpose the thin crystalline plate is placed at S. In this case a cone of light is incident on the plate, and the various rays of this cone are variously inclined to it, so that their components suffer different alterations of phase in traversing it, and therefore depict various colours on the screen F″F″. Consequently curved fringes will be seen on the screen, the form of which will depend on the position of the axes of the crystalline plate with respect to its faces.

For the success of these experiments it is necessary to work with strong light, on account of the magnitude of the image on the screen. The solar light may be replaced by electric light, cast by a lens on the polarising Nicol in a parallel beam.

### 2. *Convergent or Divergent Plane-Polarised Light*

231. By placing the crystalline plate between the lenses L and L′ (Fig. 178) the phenomena produced by converging or diverging light may be studied. It is clear that in this case, if with the point O (Fig. 179) (where the axis of the incident cone of light meets the crystal) as centre, we describe any circle on the face of the plate, all the rays of the conical shell which fall upon this circle will meet the plate at the same angle of incidence. At each point of the plate the intensity will have a definite value depending on the value of the retardation $\delta$, which in

turn depends on the angles of incidence and refraction, as shown in the example of p. 99.

We shall now consider the important and striking phenomena which are presented when a converging or diverging pencil of plane-polarised light is transmitted through a thin plate of a uniaxal crystal, the faces of the plate being perpendicular to the optic axis.

**232. Uniaxal Crystal—Plate cut at Right Angles to Axis.**—Let a plate of a uniaxal crystal, cut perpendicularly to the axis, be placed in a conical pencil of light so that the axis of the cone is perpendicular to the face of the plate, and consequently passes through it in the direction of the optic axis.

Let the axis of the conical pencil meet the plate at O (Fig. 179), and let X (Fig. 180) be any point on the face of the crystal supposed parallel to the plane of the paper. Let OP and OA be drawn parallel

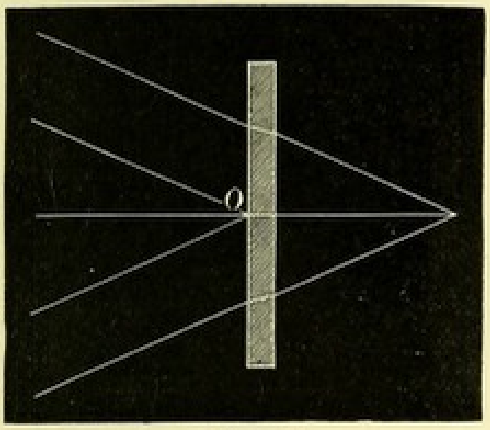



Fig. 179.

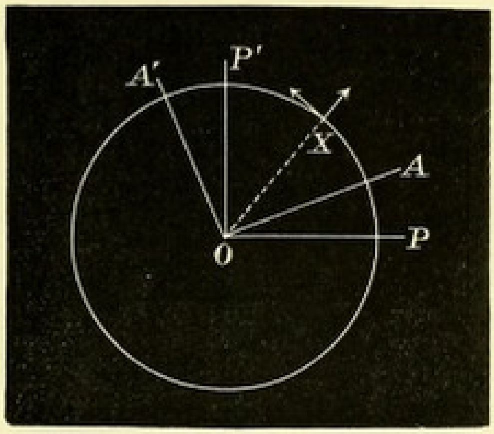



Fig. 180.

to the principal planes of the polariser and analyser. Then the plane of incidence of the ray at the point X is a plane through OX perpendicular to the face of the crystal, and this is the principal plane of the crystal at X since the optic axis is a normal to the face, for the principal plane by definition is the plane containing the ray and the optic axis. Hence a vibration incident at X, or at any point along the line OX, is split up into two components, one vibrating parallel to OX—that is, in the principal plane—and the other perpendicular to it. Consequently, if OX makes an angle $\alpha$ with OP and an angle $\beta$ with OA, the intensity of the illumination emerging from the plate at X is, after analysation (Art. 225),

$$I = a^2\{\cos^2(\alpha - \beta) - \sin 2\alpha \sin 2\beta \sin^2 \tfrac{1}{2}\delta\},$$

for homogeneous light.

If we denote the variable angle XOP by $\theta$, and the angle AOP by $\gamma$—that is, if we write $\theta$ for $\alpha$ and $\gamma$ for $\alpha - \beta$ the expression for the intensity becomes

$$I = a^2\{\cos^2 \gamma - \sin 2\theta \sin 2(\theta - \gamma) \sin^2 \tfrac{1}{2}\delta\}.$$

Now if the point X be anywhere on the line OP or on its perpendicular OP′, the incident vibration, being parallel to OP, will be in the principal plane in the former case and perpendicular to it in the latter, so that it will pass through the plate without decomposition until it is finally resolved in the analyser. The lines OP and OP′ should therefore be each of uniform illumination, and should exhibit no colour. Our formula points to the same conclusion, for if X is on OP or OP′ we have $\alpha = 0^\circ$ or $90^\circ$, the term on which the colour depends vanishes, and in both cases

$$I = a^2 \cos^2(\alpha - \beta) = a^2 \cos^2 \gamma,$$

where $\gamma$ is the angle between the principal planes of the polariser and analyser. We have therefore a rectangular cross of uniform illumination $a^2 \cos^2 \gamma$, having one arm parallel to and the other perpendicular to the principal plane of the polariser.

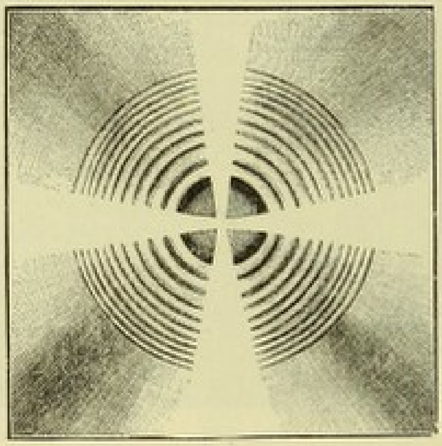


Fig. 181.—($\alpha - \beta = 0^\circ$).

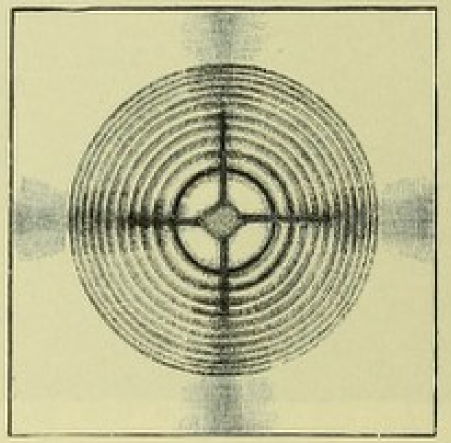


Fig. 182.—($\alpha - \beta = 90^\circ$).

Similarly, if $\beta = 0^\circ$ or $90^\circ$, we have the same value of I, viz. $I = a^2 \cos^2 \gamma$, and hence another uniform cross exists in the field having its arms respectively parallel and perpendicular to the principal plane of the analyser.

Crosses. In general, then, two rectangular crosses are seen in the field, and these are of the same uniform illumination, $a^2 \cos^2 \gamma$, and uncoloured.

If $\gamma = 0^\circ$—that is, if the principal planes of the polariser and analyser are parallel—the two crosses coincide, and we have one white cross of uniform illumination $a^2$, viz. that of the incident light (Fig. 181).

If $\gamma = 90^\circ$—that is, if the Nicols are crossed—the two crosses again coincide. But in this case the single cross is black, for $\cos \gamma = 0$ and the intensity is zero (Fig. 182).

Rings. Again, if with O as centre, we describe any circle, it is clear that the rays incident at the various points of this circle make the same angle with the normal to the plate—that is, with the optic axis. It is obvious, therefore, that the difference of phase $\delta$ introduced by the plate is the same at all points of this circle, and hence with white light there will be a series of coloured circles in the field concentric with O, and on these circles the crosses or brushes already mentioned will be superposed.

If $\sin 2\alpha \sin 2\beta$ is positive, the points of maximum intensity along a given radius between the brushes are determined by

$$\sin^2 \tfrac{1}{2}\delta = 0, \quad \text{or } \delta = 2n\pi,$$

and the minima by

$$\sin^2 \tfrac{1}{2}\delta = 1, \quad \text{or } \delta = (2n+1)\pi.$$

But if $\sin 2\alpha \sin 2\beta$ be negative the intensity will be least when $\delta = 2n\pi$, and greatest when $\delta = (2n+1)\pi$.

The curves of equal intensity for a given position of the Nicols are determined by the equation

$$\cos^2 \gamma - \sin 2\theta \sin 2(\theta - \gamma) \sin^2 \tfrac{1}{2}\delta = \text{constant}.$$

The effect of superposing on a uniaxal plate, cut perpendicularly to the axis, another plate cut in a similar manner is the same as an increase or a decrease in the thickness of the first plate according as they are of the same or of opposite signs. If they are of the same sign the rings contract, and if they are opposite signs the rings dilate. We are thus afforded with a ready and simple means of determining the sign of a crystal, by comparison with another crystal of known sign. Sign test.

**233. Isochromatic Lines and Achromatic Lines.**—When a beam of divergent or convergent polarised light is transmitted through a thin crystalline plate and examined by means of an analyser, the field is in general traversed by two systems of lines. If white light is used one of the systems of lines is brilliantly coloured, and these are termed the *isochromatic lines* or simply *the fringes*. Their form depends on the nature of the section of the crystal—that is, on the direction in which it has been cut with reference to the optic axis—and their brightness depends on the position of the analyser, the colour being most distinctly marked when the polariser and analyser are crossed ($\alpha - \beta = 90°$), as indicated by the equation

$$I = a^2\{\cos^2(\alpha - \beta) - \sin 2\alpha \sin 2\beta \sin^2 \tfrac{1}{2}\delta\}.$$

The lines of the second system are not coloured. They intersect the fringes and are termed the *achromatic* or *neutral lines*, and, as the above equation shows, they are determined by the equation

$$\sin 2\alpha \sin 2\beta = 0,$$

that is, $\alpha = 0^\circ$ or $90^\circ$, and $\beta = 0^\circ$ or $90^\circ$, while the points of maximum intensity are determined by

$$\sin^2 \tfrac{1}{2}\delta = 0, \quad \text{or } \delta = 2n\pi,$$

and the minimum points by

$$\sin^2 \tfrac{1}{2}\delta = 1, \quad \text{or } \delta = (2n+1)\pi,$$

if the product $\sin 2\alpha \sin 2\beta$ be positive; but if this produce be negative the maximum points will be determined by

$$\sin^2 \tfrac{1}{2}\delta = 1, \quad \text{or } \delta = (2n+1)\pi,$$

and the minimum points by

$$\sin^2 \tfrac{1}{2}\delta = 0, \quad \text{or } \delta = 2n\pi.$$

The maximum lines in one case, therefore, correspond to the minimum lines in the other.

Change of tint.

Now it is easily seen that $\sin 2\alpha \sin 2\beta$ generally changes sign in passing across a neutral line. For such a line arises from $\sin 2\alpha = 0$, or $\sin 2\beta = 0$, or both; but when any function passes through a zero value its sign in general changes. Thus $\sin 2\alpha$ passes through zero and changes sign as $\alpha$ passes through the value $\frac{1}{2}\pi$. Hence in crossing a neutral line corresponding to $\sin 2\alpha = 0$ the quantity $\sin 2\alpha \sin 2\beta$ will change from a positive to a negative value or *vice versâ*, and consequently the bright fringes at one side of the neutral line will correspond to the dark bands at the other. If, however, $\sin 2\alpha = 0$ and $\sin 2\beta = 0$ simultaneously, the quantity $\sin 2\alpha \sin 2\beta$ will not change sign, and the fringes will preserve the same tint in crossing the line.

Hence in crossing a neutral line the isochromatic fringes in general change to the complementary tint, except when $\sin 2\alpha$ and $\sin 2\beta$ vanish simultaneously; an example of this latter case occurs when we have a uniaxal plate cut perpendicularly to the axis and $\alpha - \beta = 0^\circ$ or $90^\circ$. Here the neutral line is a rectangular cross; $\sin 2\alpha$ and $\sin 2\beta$ vanish simultaneously, and the rings in each quadrant are the same as those in the adjacent quadrants.

We shall now consider the isochromatic lines, or fringes, which are found to be as follows—

(1) In uniaxal crystals the fringes are—

Circles for a section perpendicular to the axis.
Hyperbolas for a section parallel to the axis.
Elliptic or hyperbolic arcs for an oblique section.

(2) In biaxal crystals they are—

Closed rings for a section perpendicular to an axis.

Hyperbolas for a section parallel to the plane of the axes.

Lemniscates for a section perpendicular to the bisector of the angle between the axes.

**234. Isochromatic Surface in Uniaxal Crystals.**—The fringes presented in the case of a uniaxal crystal cut perpendicularly to the optic axis have been already discussed (Art. 232, 233). We shall now investigate the general case, in which the thin plate may be cut from the crystal in any direction, following M. Bertin.[1]

The characteristic of a fringe is that the retardation $\delta$ is the same at every point of it, and the locus of points in space for which $\delta$ has a given constant value will be an *isochromatic surface*. To every value of $\delta$ there will be a corresponding surface, and if with the radiant point as origin a system of these surfaces be described corresponding to retardations of 1, 2, 3, 4, etc., half-wave lengths, their intersections with the face of the crystal will determine the corresponding isochromatic curves or fringes exhibited in the field of view.

Let S (Fig. 183) be the radiant point, and SL any ray falling upon the face of the crystal. In general SL will be divided into two rays within the crystal, travelling with different velocities, but as the crystal is thin we may, in the approximate calculation, suppose both rays to travel along the same line LM, one with the ordinary velocity $v_o$ and the other with a velocity $v$.

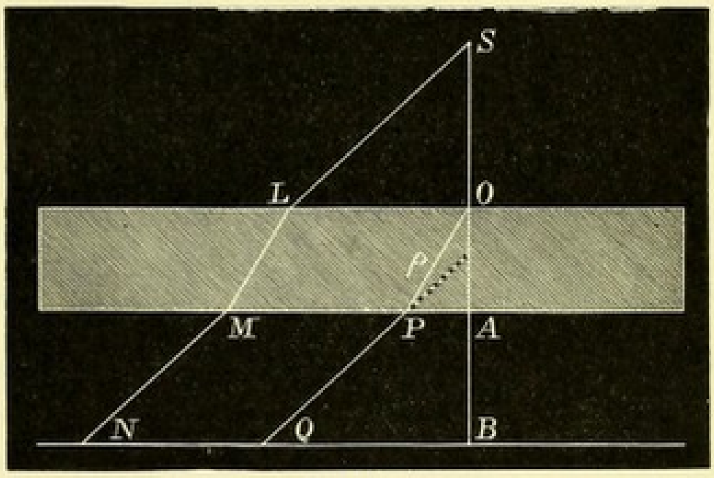


Fig 183.

The times occupied in traversing the distance LM will be $t_o = \mathrm{LM}/v_o$ and $t = \mathrm{LM}/v$ respectively, so that the time retardation is

$$t_o - t = \mathrm{LM}\left(\frac{1}{v_o} - \frac{1}{v}\right),$$

and the path retardation is

$$\delta = (\mu_o - \mu)\mathrm{LM}.$$

For the calculation of the isochromatic surfaces we may take the

[1] M. Bertin, "Mémoire sur la Surface Isochromatique," *Annales de Chimie et de Physique*, third series, tom. lxiii. p. 57, 1861.

origin at O, the point where the axis of the conical pencil meets the plate, and if we draw OP parallel to LM the retardation along OP will be the same as that along LM.

Now the wave-surface in a uniaxal crystal consists of a sphere of radius $b$ or $1/\mu_o$ (in a negative crystal like Iceland spar), and a spheroid of which the generating curve is

$$\mu_o^2 x^2 + \mu_e^2 y^2 = 1 \qquad (1).$$

If $r$ denotes any radius vector of this ellipse we have $\mu$ inversely proportional to $r$, and hence for the retardation in traversing a thickness $\rho$ we have

$$\delta = \rho(\mu_o - \mu) \qquad (2).$$

But by (1)

$$\mu^2 = \mu_o^2 \cos^2\theta + \mu_e^2 \sin^2\theta \qquad (3),$$

and by (2)

$$\mu^2 = \left(\frac{\delta}{\rho} - \mu_o\right)^2 \qquad (4).$$

Therefore, by combining (3) and (4) we find

$$\left(\frac{\delta}{\rho} - \mu_o\right)^2 = \mu_o^2 \cos^2\theta + \mu_e^2 \sin^2\theta.$$

Hence

$$(\delta - \rho\mu_o)^2 = \mu_o^2 x^2 + \mu_e^2 y^2,$$

which, since $\rho^2 = x^2 + y^2$, gives

$$\{(\mu_e^2 - \mu_o^2)y^2 - \delta^2\}^2 = 4\mu_o^2\delta^2(x^2 + y^2),$$

the generating curve of the isochromatic surface.

The isochromatic surface is formed by the revolution of this curve round the axis of the crystal. Its general form is represented in Fig. 184. In the neighbourhood of the axis of $y$ the curve resembles an hyperbola, and the surface an hyperboloid of revolution. By assigning various values to $\delta$ we determine corresponding surfaces, and the intersections of these with the face of the crystal give the isochromatic curves or fringes. Thus if $\delta = \pi$, which corresponds to a retardation $\frac{1}{2}\lambda$, the corresponding surface intersects the face of the crystal in the first dark fringe, etc.

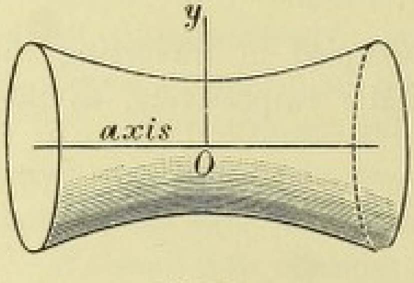



Fig. 184.

(1) *Section Perpendicular to the Optic Axis.*—If the thickness of the plate be $e$ the section of the isochromatic surface by the face of the

plate cut perpendicular to the axis is found by putting $x = e$. The curve is obviously a circle of which the radius is $r = y$ determined by the equation

$$(\mu_e^2 - \mu_o^2)^2 r^4 - 2\delta^2(\mu_e^2 + \mu_o^2)r^2 - 4\mu_o^2\delta^2 e^2 + \delta^4 = 0,$$

or

$$r^2 = \frac{\delta^2(\mu_e^2 + \mu_o^2) \pm 2\mu_o\delta\sqrt{\mu_e^2\delta^2 + (\mu_e^2 - \mu_o^2)^2 e^2}}{(\mu_e^2 - \mu_o^2)^2}.$$

Since $\delta$ is small compared with $e$, the expression for $r$ reduces to the approximate form

$$r^2 = \frac{2\mu_o\delta e}{\mu_e^2 - \mu_o^2},$$

or since $\mu_e + \mu_o = 2\mu_o$ nearly, we have the approximate expression

$$r^2 = \frac{\delta e}{\mu_e - \mu_o}.$$

The consecutive rings are determined by making $n$ equal to the consecutive whole numbers in the equation $\delta = n\frac{1}{2}\lambda$, the dark rings corresponding to the odd values of $n$ and the bright rings to the even values.

The above expression for the radii of the rings has been obtained on the supposition that the radiant point is at O on the surface of the crystal. If the luminous origin be at S the radius of the ring on the face of the plate is $R = AM$, but in the above calculation $r = AP$, therefore

$$\frac{R}{r} = \frac{AM}{AP} = \frac{D\tan i}{e\tan r} = \frac{\mu_o D}{e},$$

if $i$ and $r$ be small and D denotes the distance of S from the plate.

Hence the radii of the rings on the plate are given by the equation

$$R^2 = \frac{2\mu_o^3 D^2\delta}{e(\mu_o^2 - \mu_e^2)} = \frac{\mu_o^2 D^2\delta}{e(\mu_o - \mu_e)}\ q.p.,$$

where for the bright rings $\delta = 2n\frac{1}{2}\lambda$ and for the dark rings $\delta = (2n+1)\frac{1}{2}\lambda$.

If the rings be received on a screen at a distance $D'$ from the plate, the radii $R'$ will be found from the equation

$$\frac{R'}{R} = \frac{BN}{AM} = \frac{D + D'}{D}.$$

If we regard $\mu_o$ and $\mu_e$ as approximate constants we see that the radii of the rings are directly proportional to the square root of the wave length $(\delta = n\frac{1}{2}\lambda)$ and inversely as the square root of the thickness of the plate. The diameters of the consecutive rings are also proportional to the square roots of the numbers 1, 2, 3, 4, etc.

(2) *Section Parallel to the Axis.*—[As the equation of the surface is

$$\{(\mu_e^2 - \mu_o^2)(y^2 + z^2) - \delta^2\}^2 = 4\mu_o^2\delta^2(x^2 + y^2 + z^2),$$

we obtain the equation of a section parallel to the plane of $x$ and $y$ by putting $z=e$. Close to the axis of $z$ this curve approximates to the hyperbola

$$2y^2(\mu_e{}^2-\mu_o{}^2)\{(\mu_e{}^2-\mu_o{}^2)e^2-\delta^2\}-4(x^2+y^2)\mu_o{}^2\delta^2+\{(\mu_e{}^2-\mu_o{}^2)e^2-\delta^2\}^2-4\mu_o{}^2\delta^2e^2=0.$$

Its equation is found by neglecting the fourth power of $y$ after replacing $z$ by $e$. Corresponding to various values of $\delta$ we have a system of hyperbolæ, and the figure represents the curves corresponding to $\delta=\frac{1}{2}n\lambda$, $n$ being any whole number. If the absolute term vanishes the corresponding hyperbolæ reduce to a pair of lines, and the appropriate value of $\delta$ is easily seen to be $\delta=(\mu_e-\mu_o)e$, for $\delta$ must be small in comparison with $e$, and the solution $\delta=(\mu_e+\mu_o)e$ must consequently be rejected. On substitution for $\delta$ we find the equation of the lines to be

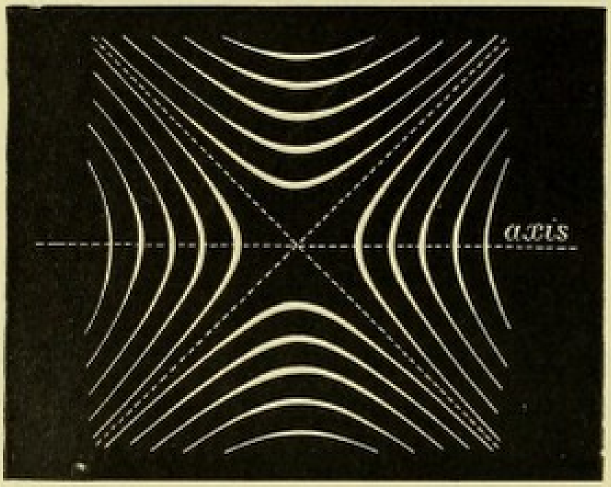



Fig. 185.

$$\mu_e y^2-\mu_o x^2=0.$$

and as $\mu_e$ is nearly equal to $\mu_o$ the lines are very close to the bisectors of the angles between the axis of $x$ and $y$, and are therefore nearly at right angles. If $(\mu_e-\mu_o)e$ is equal to a whole number of half-wave lengths, these lines are comprised in the system of fringes.

In order to obtain a clearer idea of the arrangement of the hyperbolæ we replace $\delta$ by $f(\mu_e-\mu_o)e$ in the equation, and after dividing across by $(\mu_e-\mu_o)^2$ we put $\mu_e=\mu_o$, and so obtain for the equation of the system

$$y^2(2-f^2)-x^2f^2+e^2(1-f^2)=0.$$

If $f$ is less than unity—that is, if the retardation is less than $(\mu_e-\mu_o)e$—the hyperbola cuts the axis of $x$. If $f>1$, it cuts the axis of $y$, and the lines corresponding to $f=1$ separate the two sets of curves.] The fringes are most brilliant when the polariser and analyser are parallel, or crossed, and make an angle of 45° with the principal plane of the plate, but when the principal plane is parallel or perpendicular to the principal section of the analyser the fringes disappear entirely.

(3) *Oblique Section.*—The section of an isochromatic surface by a plane inclined to the axis is a curve of the fourth degree which approximates to a circle when the plane is nearly perpendicular to the axis

and to an hyperbola when the plane is nearly parallel to the axis. For an obliquity of 45° it is only the portions of the fringes near their vertices that are seen, and these are right lines perpendicular to the principal section of the plate.

**235. Isochromatic Surface in Biaxal Crystals.**—In the case of biaxal crystals the relative retardation of the rays in traversing a distance $\rho$ may, as before, be written in the form

$$\delta = \rho(\mu' - \mu'')$$

where $\mu'$ and $\mu''$ are the roots of the equation

$$\frac{l^2}{\mu^2 - \mu_1^2} + \frac{m^2}{\mu^2 - \mu_2^2} + \frac{n^2}{\mu^2 - \mu_3^2} = 0,$$

in which $l$, $m$, $n$ are the direction cosines of $\rho$, and $\rho$ is a radius vector of the isochromatic surface when $\delta$ is constant (see Art. 205).

Now from this equation we have

$$\mu'^2 + \mu''^2 = \Sigma(\mu_2^2 + \mu_3^2)l^2,$$

and

$$\mu'^2\mu''^2 = \Sigma\mu_2^2\mu_3^2 l^2.$$

But since

$$\frac{\delta}{\rho} = \mu' - \mu'',$$

we have

$$\left(\mu'^2 + \mu''^2 - \frac{\delta^2}{\rho^2}\right)^2 = 4\mu'^2\mu''^2.$$

Hence

$$\left\{\Sigma(\mu_2^2 + \mu_3^2)l^2 - \frac{\delta^2}{\rho^2}\right\}^2 = 4\Sigma\mu_2^2\mu_3^2 l^2,$$

or in Cartesian co-ordinates, the equation of the isochromatic surface is

$$\{(\mu_2^2 + \mu_3^2)x^2 + (\mu_3^2 + \mu_1^2)y^2 + (\mu_1^2 + \mu_2^2)z^2 - \delta^2\}^2$$
$$= 4(x^2 + y^2 + z^2)(\mu_2^2\mu_3^2x^2 + \mu_3^2\mu_1^2y^2 + \mu_1^2\mu_2^2z^2).$$

All the surfaces are obtained by giving various values to $\delta$. They are similarly situated, and their intersections with the face of the crystal give the isochromatic curves, each corresponding to a certain constant difference of phase.

The form of these surfaces is represented[1] in Fig. 186. The section by the plane $xz$ which contains the optic axes is found by

[1] M. Bertin remarks that "la surface isochromatique des cristaux à deux axes ressemble à une croix de Saint-André dont les bras seraient cylindriques et dirigés suivant les axes optiques du cristal."

making $y$ zero in the above equation. It consists of two branches, one having asymptotes parallel to the optic axes, the other interior to it is closed and parasitic, not being related to the question in hand. It is obvious that in the directions of the optic axes the values of $\rho$ should be infinite, for since these are the optic axes $v'=v''$ or $\mu'=\mu''$, and therefore $\rho$ is infinite for a given retardation.

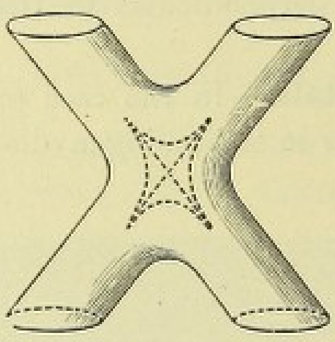


Fig. 186.

The sections made by the other co-ordinate planes are closed curves of the fourth degree.

(1) *Section Parallel to the Plane of the Axes.*—The sections of the isochromatic surface by a plane parallel to that containing the optic axes differ little near the centre from hyperbolas. [The equation of these hyperbolas is found by replacing $y$ by $e$ and neglecting powers and products of the fourth order in $x$ and $z$. By a process similar to that employed in section (2) of the last article we find

$$x^2\left\{2\frac{\mu_3-\mu_2}{\mu_3-\mu_1}-f^2\right\}+z^2\left\{2\frac{\mu_2-\mu_1}{\mu_3-\mu_1}-f^2\right\}+e^2(1-f^2)=0,$$

where the retardation is given by $\delta=ef(\mu_3-\mu_1)$. If $f=1$, the hyperbola nearly coincides with the lines $x^2-z^2=0$, or more exactly with the optic axes $\mu_3{}^2x^2-\mu_1{}^2z^2=0$.] These fringes are like those of a uniaxal crystal cut parallel to its axis (Fig. 185).

(2) *Section Perpendicular to a Bisector of the Angle between the Optic Axes.*—The sections of the isochromatic surfaces by a plane perpendicular to the internal or external bisector of the angle between the

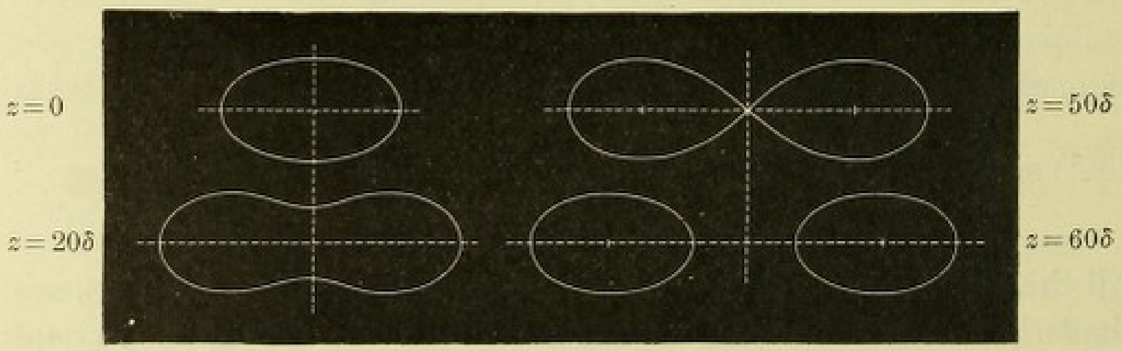



Fig. 187.

optic axes—that is, parallel to $xy$—resemble the several forms of lemniscates.

The sections of the surface corresponding to a given value of $\delta$ by planes drawn at right angles to the plane of the optic axes, and at distances $z=20\delta$, $z=50\delta$, $z=60\delta$ respectively from the centre, are represented in Fig. 187.

In this case we can determine the approximate form of the fringes directly. For if OM and OM′ (Fig. 188) be the optic axes, $\mu'$ and $\mu''$ the indices for the rays travelling in any direction OP which makes angles $\theta$ and $\theta'$ with the axes, then

$$\delta = \mathrm{OP}(\mu' - \mu'').$$

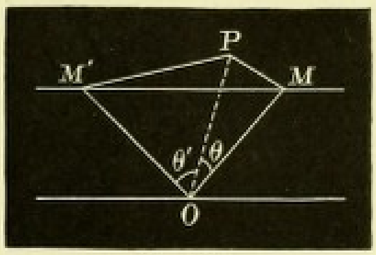


Fig. 188.

But by Art. 206, Cor. 1, we have the approximate formula

$$\mu' - \mu'' = (\mu_1 - \mu_3) \sin\theta \sin\theta',$$

and approximately

$$\sin\theta = \frac{\mathrm{PM}}{\mathrm{OM}}, \quad \sin\theta' = \frac{\mathrm{PM}'}{\mathrm{OM}'}, \quad \mathrm{OP} = e.$$

Therefore

$$\frac{\delta}{e} = (\mu' - \mu'') = (\mu_1 - \mu_3)\frac{\mathrm{PM}\,.\,\mathrm{PM}'}{(\mathrm{OM})^2},$$

that is,

$$\mathrm{PM}\,.\,\mathrm{PM}' = \text{constant},$$

which is the polar equation of a lemniscate having M and M′ for foci.

(3) *Section Perpendicular to an Optic Axis.*—In this case the isochromatic lines will be a system of closed curves resembling deformed lemniscates encircling the axis to which the section is perpendicular.

**236. Achromatic or Neutral Lines.**—The uncoloured lines are determined by the equation

$$\sin 2\alpha \sin 2\beta = 0.$$

They are consequently the locus of a point on the face of the crystal such that the planes of polarisation of the rays in the crystal are at that point either parallel or perpendicular to the principal plane of the polariser or analyser.

(1) *Section Perpendicular to the Bisector of the Angle between the Optic Axes.*—The planes of polarisation of the rays travelling in any direction OP (Fig. 188) are the bisectors of the angles between the planes OPM and OPM′, which contain the ray OP and the optic axes OM and OM′ respectively. The traces of these planes on the face of the crystal will, for a small angle of incidence, approximately coincide with the bisectors of the angle MPM′. Hence if the plane of the paper be taken parallel to the face of the crystal, and if the optic axes meet it at M and M′ (Fig. 189), the point P will describe a neutral line if the bisector of the angle MPM′ is parallel or perpendicular to the principal plane of the polariser or analyser. Let OX be parallel to the principal plane of the polariser, and take OX and the perpendicular line OY as axes of reference, O being the middle point of MM′.

The bisector of the angle MPM′ must be perpendicular to OX, therefore the triangle APB is isosceles. Hence if the co-ordinates of P be $x$, $y$, and those of M be $x'$, $y'$, we have

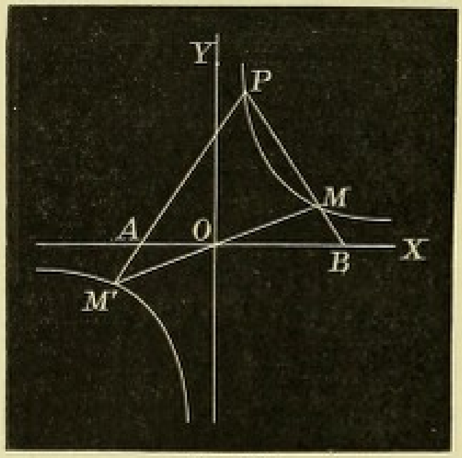


Fig. 189.

$$\cot \text{PBO} = \frac{x' - x}{y - y'}, \quad \text{and} \cot \text{PAO} = \frac{x + x'}{y + y'},$$

equating these expressions we find at once

$$xy = x'y'.$$

The locus of P is consequently a rectangular hyperbola, passing through M and M′, of which the asymptotes are OX and OY, lines parallel and perpendicular to the trace of the principal plane of the polariser.

In the same manner we find a second rectangular hyperbola corresponding to $\beta = 0^\circ$ or $90^\circ$, also passing through M and M′, and having for asymptotes lines parallel and perpendicular to the trace of the principal plane of the analyser.

The uncoloured lines therefore consist of four hyperbolic branches passing two and two through M and M′, the apparent extremities of the optic axes, the asymptotes of one curve being a parallel and a perpendicular to the trace of the principal plane of the polariser, while those of the other bear a corresponding relation to the principal plane of the analyser.

If the principal plane of the polariser is parallel to MM′—that is, if MM′ coincides with OX, we have $\beta = 0^\circ$, the hyperbola reduces to its asymptotes $xy = 0$, and becomes a rectangular cross. If in addition the optic axes coincide, the other hyperbola will also become

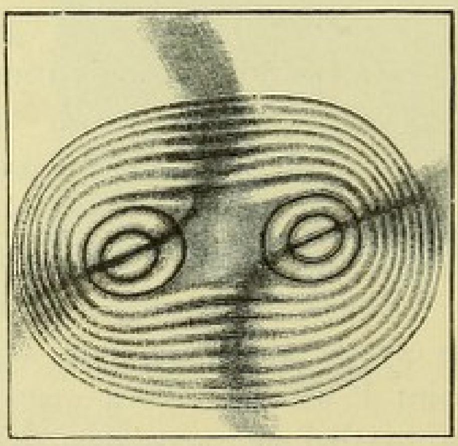


Fig. 190.

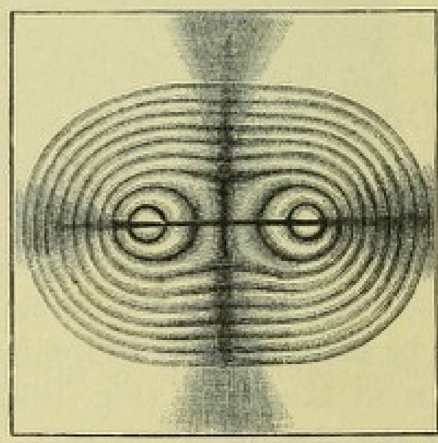


Fig. 191.

a rectangular cross, as we have already seen to be the case for a uniaxal crystal.

If the polariser and analyser are either parallel or crossed, the two hyperbolas coincide, and we have a single rectangular hyperbola of which one branch passes through M and the other through M′, as shown in Fig. 190, while if in addition the plane of polarisation be parallel or perpendicular to MM′ the curve reduces to a rectangular cross as shown in Fig. 191. When the polariser and analyser are parallel we have $\alpha - \beta = 0$, $I = a^2$, and the brush is bright, but if the Nicols be crossed $\alpha - \beta = 90^\circ$, and the brush is dark.

The neutral hyperbolic bands are made to pass through all possible forms by rotating the plate between the Nicols, and, as in the case of uniaxal crystals, the coloured fringes change to the complementary tint in crossing the neutral lines (Art. 233).

(2) *Section Parallel to the Plane of the Axes.*—In this case the optic axes OM and OM′ (Fig. 192) at any point O in the face of the crystal lie in the plane of the face, and for a ray incident nearly normally at O, the planes of polarisation of the two refracted rays will bisect the angle MOM′ internally and externally. Hence if the angle of incidence be small the bisectors OX and OY will have the same direction at every point, and therefore if they are parallel or perpendicular to the principal plane of the polariser or analyser at one point they will be so at all, and the whole field will be uncoloured. There are consequently no neutral lines, but in four positions the whole field is uncoloured.

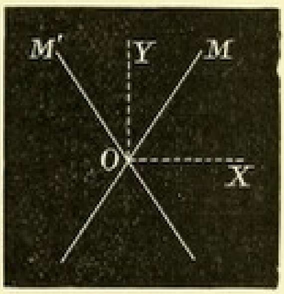


Fig. 192.

(3) *Section Perpendicular to an Axis.*—When the plate is cut at right angles to one of the optic axes the fringes form a system of concentric rings, and these are crossed by two neutral bars which in general are not rectangular. When the Nicols are parallel or crossed the two bars coincide and the neutral line is sensibly straight, and either white or black.

### 3. *Circularly and Elliptically Polarised Light*

**237. Parallel Circularly Polarised Light.**—The phenomena which are presented when circularly or elliptically polarised light is transmitted through a thin crystalline plate were investigated by Airy.[1] The characteristic of circularly polarised light is that it results from two equal plane-polarised parts, polarised at right angles and differing in phase by a quarter of a period. Its components in any

[1] Airy, *Camb. Trans.* vol. iv.

two rectangular directions may therefore be represented by the equations,

$$x = a\cos\omega t, \quad \text{and } y = a\sin\omega t,$$

the intensity of the original beam being $2a^2$. It may be obtained by transmitting plane-polarised light through a quarter-wave plate (or a Fresnel's rhomb suitably placed), the action of these instruments being to divide the incident plane ray into two others vibrating in perpendicular directions, and differing in phase by a quarter period. The determination of the phenomena which arise when circularly polarised light is transmitted through a crystalline plate may therefore be obtained from Art. 229 by introducing the condition that the first plate introduces a phase difference of a quarter period. The direct investigation of the intensity in the field of the analyser is, however, very simple. Let the components of the circular vibration falling on the plate be

$$x = a\cos\omega t, \quad \text{and } y = a\sin\omega t,$$

parallel to the two directions of vibration in the plate as in Art. 211. The passage through the plate introduces a relative difference of phase $\delta$, so that the emerging vibrations may be written in the form

$$x = a\cos\omega t, \quad \text{and } y = a\sin(\omega t + \delta).$$

These, resolved parallel to the principal plane of the analyser, give the vibration

$$a\cos\alpha\cos\omega t + a\sin\alpha\sin(\omega t + \delta) = a(\cos\alpha + \sin\alpha\sin\delta)\cos\omega t + a\sin\alpha\cos\delta\sin\omega t.$$

The intensity is therefore

$$\begin{aligned} \mathrm{I} &= a^2\{(\cos\alpha + \sin\alpha\sin\delta)^2 + (\sin\alpha\cos\delta)^2\}, \\ &= a^2(1 + \sin 2\alpha\sin\delta), \end{aligned}$$

which is independent of the position of the polariser.

If the analyser be merely a doubly refracting rhomb we shall have in the same manner for the ordinary image

$$\mathrm{I}' = a^2(1 - \sin 2\alpha\sin\delta),$$

and therefore

$$\mathrm{I} + \mathrm{I}' = 2a^2,$$

so that the images are complementary.

Had we started on the supposition that the $x$-component of the vibration is retarded relatively to the $y$-component, we would have had

$$x = a\cos(\omega t + \delta), \quad \text{and } y = a\sin\omega t,$$

and for the intensity in the field of the analyser we would have the complementary expressions

$$\mathrm{I} = a^2(1 - \sin 2\alpha \sin \delta),$$
$$\mathrm{I}' = a^2(1 + \sin 2\alpha \sin \delta).$$

Hence if the sense of the circular polarisation be reversed the two images in the field of a doubly refracting analyser will interchange tints. The formula for I shows that the image possesses a certain colour depending on the term sin $2\alpha$ sin $\delta$, and that there are two positions in which the image is achromatic, given by the equation Colour term.

$$\sin 2\alpha = 0,$$

that is, when the analyser is parallel or perpendicular to the principal section of the thin plate, and in both these positions the intensity is the same and equal to half that of the incident light, for we have

$$\mathrm{I} = \mathrm{I}' = a^2.$$

The tint, depending on $\delta$, will vary with the thickness of the plate, but here it will not follow the law of colour in Newton's rings which holds in the case of plane-polarised light, for the term on which the colour depends is here proportional to sin $\delta$ and not to $\sin^2 \frac{1}{2}\delta$.

If a quarter-wave plate or a Fresnel's rhomb be introduced between the crystal and the analysing Nicol in such a manner that light passing through the Nicol would be circularly polarised by the plate (that is, so that the principal plane of the Nicol makes an angle of 45° with the principal plane of the plate), then the light falling on the crystal will be circularly polarised and the light emerging from it will be circularly analysed. In this case the vibrations emerging from the crystal are, as before, of the form Light circularly analysed.

$$a \cos \omega t, \quad \text{and } a \sin (\omega t + \delta).$$

Hence if the principal plane of the quarter-wave plate be inclined at an angle $\alpha$ to that of the crystal the two components emerging from the plate are

$$a \cos \alpha \cos \omega t + a \sin \alpha \sin (\omega t + \delta),$$

and

$$a \sin \alpha \sin \omega t + a \cos \alpha \cos (\omega t + \delta).$$

These resolved parallel to the principal plane of the analysing Nicol, which is at 45°, give an expression which is simply their sum multiplied by $1/\sqrt{2}$—that is,

$$\frac{a}{\sqrt{2}} \left\{ \cos (\omega t - \alpha) + \cos (\omega t + \delta - \alpha) \right\}.$$

Hence

$$I=2a^2\cos^2\tfrac{1}{2}\delta.$$

There are consequently no neutral lines, and the isochromatic lines are the same as when the light is plane-polarised and analysed.

**238. Convergent Circularly Polarised Light.**—If the incident light be convergent, achromatic lines and coloured fringes are presented as in the case of plane-polarised light, but here they are of a simpler character. The uncoloured lines are determined by the equation

$$\sin 2\alpha=0,$$

therefore they are only half as numerous as in the case of plane-polarised light, for in the latter case the neutral lines are determined by the equation $\sin 2\alpha \sin 2\beta=0$.

Thus with a uniaxal plate cut perpendicularly to the optic axis we have with plane-polarised light in general two rectangular crosses of uniform intensity and neutral tint, but with circularly polarised light we have a single cross, the arms of which are respectively parallel and perpendicular to the trace of the principal section of the analyser. The intensity of the neutral lines is uniform and equal to $a^2$, being independent of the angle between the principal planes of the polariser and analyser.

For a given position of the Nicols, if $\sin 2\alpha$ is positive, the points of maximum intensity correspond to

$$\sin\delta=1, \quad \text{or } \delta=(4n+1)\tfrac{1}{2}\pi,$$

that is,

$$\delta=\frac{\pi}{2}, \quad \frac{5\pi}{2}, \quad \frac{9\pi}{2}, \text{ etc.} \qquad \text{(max.)},$$

and the points of minimum intensity to

$$\sin\delta=-1, \quad \text{or } \delta=(4n-1)\tfrac{1}{2}\pi,$$

that is,

$$\delta=\frac{3\pi}{2}, \quad \frac{7\pi}{2}, \quad \frac{11\pi}{2}, \text{ etc.} \qquad \text{(min.)}.$$

In the case of plane-polarised light we have seen that the bright and dark fringes are determined respectively by

$$\delta=2n\pi=0, \quad 2\pi, \quad 4\pi, \text{ etc.},$$

and

$$\delta=(2n+1)\pi=\pi, \quad 3\pi, \quad 5\pi, \text{ etc.}$$

Hence the bands correspond to the even multiples of $\frac{1}{2}\pi$ in the latter case and to the odd multiples in the former.

The effect of the circular polarisation has been to displace the fringes through a quarter of an order. In one pair of opposite quadrants they are pulled in towards the centre by this amount, and in the other pair they are pushed out by the same amount, for the bright rings in any quadrant correspond to the dark rings in the adjacent quadrants, each fringe changing to its complementary in crossing the neutral lines (Figs. 193, 194).

This result affords a convenient method of determining the sign of Sign test.
a crystal, for if two crystals have opposite signs then $\sin\delta$ will be positive for one and negative for the other, consequently the bright rings afforded by one will correspond to the dark rings of the other. But we have seen that the bright rings in any quadrant correspond to the dark rings in the adjacent quadrant, therefore for a given position of

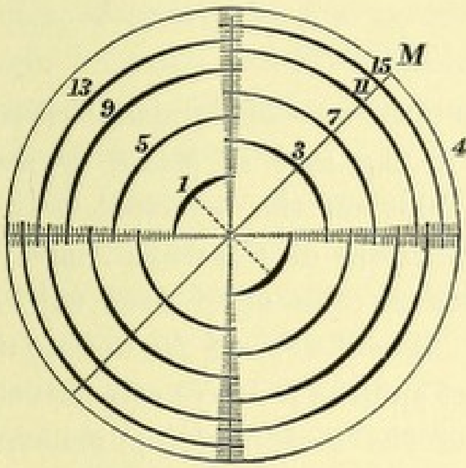



Fig. 193.

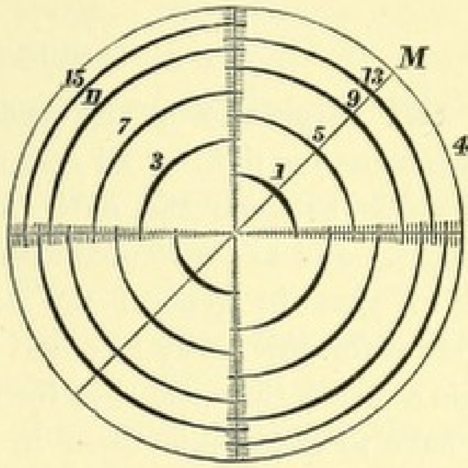



Fig. 194.

the Nicols the rings afforded by one crystal will be similar to those of the other turned through a right angle.

If the thin plate be cut from a biaxal crystal perpendicular to the mean line, the isochromatic curves are lemniscates, and crossing these we have the two branches of a rectangular hyperbola passing through the apparent extremities of the optic axes, and having for asymptotes a parallel and a perpendicular to the trace of the principal plane of the analyser. The fringes change to the complementary tint where they cross the neutral lines—that is, the bright bands in any quadrant correspond to the dark bands in the neighbouring quadrants.

**239. Elliptically Polarised Light.**—If the light incident on the plate be elliptically polarised, its components parallel to the principal directions in the plate will differ in amplitude and phase. The effect of the plate is to alter this difference of phase, so that emerging from the plate the vibrations may be written in the form

$$x = a\cos\omega t, \quad \text{and } y = b\cos(\omega t + \delta).$$

Resolving these parallel to the principal plane of the analyser we have in it the vibration

$$a \cos \alpha \cos \omega t + b \sin \alpha \cos (\omega t + \delta) = (a \cos \alpha + b \sin \alpha \cos \delta) \cos \omega t - b \sin \alpha \sin \delta \sin \omega t.$$

The intensity is consequently given by

$$\begin{aligned} \mathrm{I} &= (a \cos \alpha + b \sin \alpha \cos \delta)^2 + b^2 \sin^2 \alpha \sin^2 \delta, \\ &= a^2 \cos^2 \alpha + b^2 \sin^2 \alpha + ab \sin 2\alpha \cos \delta. \end{aligned}$$

The uncoloured lines are therefore again given by the equation

$$\sin 2\alpha = 0.$$

The forms of the fringes have been investigated by Airy in some of the simpler cases, particularly that of a uniaxal crystal cut perpendicularly to the optic axis.

### 4. *Dispersion of the Optic Axes*

**240. Dispersion of the Axes.**—The optic axis of a uniaxal crystal is the same for light of all colours, but in the case of biaxal crystals the angle between the optic axes depends on the principal indices $\mu_1$, $\mu_2$, $\mu_3$, and as these quantities are functions of the wave length, it follows that in general the angle between the optic axes will be different for differently coloured lights. Thus if $2\theta_v$ and $2\theta_r$ denote the angle between the optic axes for the violet and the red rays respectively, we have to determine them in terms of the corresponding principal velocities (Art. 201)

$$\tan \theta_v = \sqrt{\frac{a_v^2 - b_v^2}{b_v^2 - c_v^2}}, \quad \tan \theta_r = \sqrt{\frac{a_r^2 - b_r^2}{b_r^2 - c_r^2}},$$

and as $a$, $b$, $c$ do not, in general, vary proportionally in passing from one colour to another, it follows that $\theta$ will vary with the wave length. This *dispersion* of the optic axes, as it is termed, depends on the character of the crystal, presenting different peculiarities according as the crystalline axes are rectangular or oblique.

(1) *Orthorhombic System.*—Crystals of the orthorhombic (or trimetric) system possess three rectangular planes of symmetry. The crystallographic axes are unequal but mutually rectangular.

Now the optic axes are situated in the plane ($xz$) of the greatest and least axes of elasticity. Consequently if we have $a > b > c$ for all wave lengths the optic axes for all colours will be situated in the same plane ($ac$), but if it should happen that for some colours we have $a > b > c$ and for others $b > a > c$, the optic axes of the former will lie in the plane ($ac$), and those of the latter in the perpendicular plane ($bc$). So again if for any colour we have $a>c>b$, the corresponding optic axes will be situated in the plane $(ab)$. The pairs of optic axes situated in any plane have, however, a common bisector of the angle between them, viz. the axis of elasticity which lies in that plane; they are therefore symmetrically situated with respect to this line. When the variations of $a$, $b$, $c$ from colour to colour are small, compared with the differences of the quantities themselves, then the optic axes of the various colours (as in the case of aragonite) will all lie in the same plane, and be symmetrically situated with respect to the same line.

Thus the lemniscate fringes presented by a plate cut perpendicularly to the greatest or least axes of elasticity—that is, perpendicularly to either bisector of the angle between the optic axes—will possess the same centre but will have different foci for the different colours, the distance between them increasing or decreasing from the red to the violet, according to the nature of the crystal; but in some cases the angle between the axes attains a maximum value for some colours between the red and violet, and then decreases from this towards both ends of the spectrum.

(2) *Monoclinic System.*—In the monoclinic system of crystals two of the crystallographic axes are inclined at an oblique angle, while the third is perpendicular to their plane. Thus of the angles between the axes two are right and one oblique. There is consequently only a single plane of symmetry, viz. that containing the oblique axes. The third axis, which is perpendicular to this plane, preserves the same direction for all colours, but the directions of the other two axes may vary in the plane of symmetry, so that great complication may occur in the fringes. Two distinct cases arise:—

($\alpha$) If for any colour the bisector of the angle between the optic axes coincides with that crystallographic axis which is perpendicular to the plane of the other two, this line will bisect the angle between the optic axes for all the other colours, but since the oblique axes in the plane of symmetry vary with the wave length, it follows that the plane containing the optic axes may be situated in any azimuth. Borax presents this mode of dispersion, the optic axes for the various colours being situated in different planes but possessing a common bisector. Thus the isochromatic fringes afforded by a plate cut perpendicularly to the bisector of the acute angle between the optic axes (the greatest axis of elasticity in negative, and the least in positive crystals) possess the same centre, but their axes may be in any direction.

($\beta$) If the optic axes for any colour are in the plane of symmetry—that is, the plane containing the oblique axes of the crystal—then the optic axes for all colours will be situated in the same plane, but as their directions may vary in any manner, the angles between them will not have a common bisector and the isochromatic curves will not have a common centre, but their centres will be situated on a right line with respect to which the curves are symmetrical. This species of dispersion is exhibited in gypsum.

(3) *Triclinic System.*—In the triclinic system of crystals the three crystallographic axes are inclined at oblique angles, and the position of the axes of optical elasticity may vary in any manner from colour to colour. We may consequently be presented simultaneously with the modes ($\alpha$) and ($\beta$)—that is, the optic axes for the different colours may be situated in different planes and at the same time they may not have a common bisector.

*Effect of a Change of Temperature.*—An elevation of temperature generally diminishes the refractive index of a transparent substance. Hence in general a variation in the temperature of a crystal causes a corresponding variation in the quantities $a$, $b$, $c$, so that the angle between the optic axes for any given colour will in general change with the temperature of the crystal. And further, if at one temperature the optic axes for all colours lie in the same plane, then at another temperature some may remain in that plane while others are displaced to the perpendicular plane. This was observed by Brewster[1] in sulphate of sodium. In this substance the optic axes at ordinary temperatures are all situated in the same plane, but they exhibit great dispersion and the fringes are greatly confused. Operating with monochromatic light of various colours, the angle between the optic axes was found to increase from the violet to the red, being small for the former and considerable for the latter. On gradually raising the temperature the angle was seen to diminish for all the colours, the violet axes shrinking closer and closer till they finally coalesced, and the crystal for these rays then behaved as if uniaxal. On continuing the elevation of temperature the violet axes separated again, but now in the perpendicular plane, and at 60° the angle between them attained a considerable magnitude. In operating with white light the phenomena become extremely confused and indistinct.

In the case of the oblique systems of crystals an elevation of temperature may displace one axis notably more than the other, so that the mean line corresponding to this colour becomes displaced in direction.

[1] Brewster, *Phil. Mag.* (3), vol. i. p. 417.

# CHAPTER XVI

## ON THE STUDY OF POLARISED LIGHT

**241. Detection of Polarised Light.**—The fringes and colours exhibited when polarised light is transmitted through a thin crystalline plate placed before an analyser afford a most delicate test of the presence of polarisation in a beam of light, or of doubly refracting structure in a transparent substance. It is to be expected, therefore, that instruments depending in principle upon the exhibition of these fringes should have been early invented and applied to the study of polarised light. In general any instrument which acts as a polariser may also be used as an analyser. Thus a Nicol's prism may be used either to produce a pencil of plane-polarised light or to detect polarisation in any other pencil. If light be transmitted through a Nicol the intensity of the transmitted beam will vary as the Nicol is rotated if the incident beam is polarised or partially polarised. But this method of detecting the presence of polarisation depends on the variation of the intensity of an image, which will be very minute, and consequently escape observation if the quantity of polarised light in the incident beam is small.

Instruments depending on the production of colours or fringes are much more delicate, and detect the presence of small traces of polarisation. We shall describe those of Savart and Babinet.

**242. Savart's Polariscope.**—Savart's polariscope is a simple and delicate contrivance for the detection of plane-polarised light.

A thin plate of a uniaxal crystal is divided into two halves in order to secure two similar thin plates of equal thickness. The two portions are now superposed so as to form a plate of double the thickness, and one of them is rotated through 90°, so that their principal sections are at right angles. The plates, so placed, are mounted in a small tube before a Nicol's prism, or any other analyser, of which the principal section is turned parallel to the bisector of the angle between the principal sections of the plates.

When a beam of convergent or divergent plane-polarised light is allowed to fall upon the plates the field of the analyser is crossed by a system of coloured rectilinear fringes parallel to the bisector of the angle between the principal planes of the plates. The sensibility of the instrument increases as the plane of polarisation of the incident light approaches the direction of the fringes—that is, the bisector of the angle between the principal sections—and it is sufficiently delicate to easily detect the polarisation of the light of the sky.

The quartz plates are cut at an angle of about 45° to the axis in order to avoid the effects of the rotatory power which the crystal possesses in the direction of its axis (chap. xvii.).

**243. Babinet's Compensator.**—The compensator devised by M. Babinet admits of the study of polarised light by means of coloured fringes, even though the incident light be parallel. It is an exceedingly sensitive polariscope, and has been successfully adapted, in a modified form, by M. Jamin to the study of elliptically polarised light.

It consists of two slender right-angled prisms or wedges of quartz ABD and BCD (Fig. 195), placed together with their hypothenuses in contact so as to form a thin plate of rectangular cross section ABCD. In the prism ABD the optic axis is parallel to the face through AB, and to the plane of the section ABCD, while in the prism BCD the axis is parallel to the face through CD, but perpendicular to the plane of the section. Thus in both prisms the axes are parallel to the faces of the plate, but perpendicular to each other, as in Wollaston's prism (Art. 186), but on account of the extreme smallness of the angles of the wedges in this apparatus the separation of the rays during transmission is negligible.

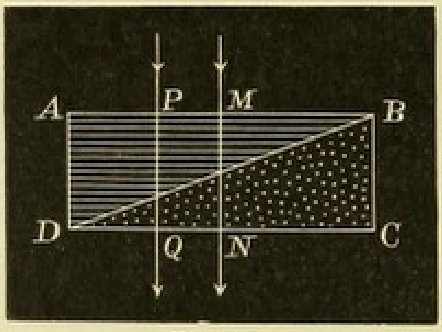


Fig. 195.

Hence if plane-polarised light falls normally on the face AB it will be broken up into two parts, one vibrating parallel to AB, and the other perpendicular to it, and these vibrations on entering the second prism will retain their directions of vibration but interchange their velocities of propagation, for the vibration parallel to AB is parallel to the optic axis in the first prism but perpendicular to it in the second. It consequently follows that the ordinary ray in the first prism becomes an extraordinary ray in the second, and *vice versâ*. Thus if any ray PQ traverses a thickness $e$ in the first prism, the relative retardation of the two vibrations will be $e(\mu_e - \mu_o)$, and for a thickness

$e'$ in the second prism will be $-e'(\mu_e - \mu_o)$, since the ray which travels fastest in the first travels slowest in the second, and *vice versâ*. Hence the whole relative path retardation of the vibrations produced by transmission through the plate is

$$\delta = (e - e')(\mu_e - \mu_o).$$

The ray MN which passes through the centre of the plate $(e = e')$ suffers no relative retardation in its component vibrations, and remains polarised in the same plane as at incidence. As we move from N towards D the component vibrations at emergence will differ in phase, but at a certain distance $a$ from N the phase difference will be equal to $\pi$, and the retardation will be $\frac{1}{2}\lambda$. At a distance $2a$ it will be $\lambda$, and so on. Consequently on the line CD we will have a system of equally spaced points at which the phase difference will be a multiple of $\pi$, and the emergent light will be plane-polarised, and on the face of the crystal there will be a corresponding system of right lines, at a common distance $a$ from each other. At the centre N, and at points on CD, distant any multiple of $2a$ from it, the phase difference will be a multiple of $2\pi$, and the transmitted light will be polarised in the same plane as the incident, but at the intermediate points, viz. those distant from N by an odd multiple of $a$, the phase difference will be an odd multiple of $\pi$, and the transmitted light will be plane-polarised also, but in this case the plane of polarisation will be inclined at an angle $2\alpha$ to the plane of polarisation of the incident light, where $\alpha$ is the angle between the primitive plane of polarisation and the principal plane of the face AB (see Art. 47, Cor. 1). The points at which the plane polarisation exists are determined by the equation

$$(e - e')(\mu_e - \mu_o) = n \,.\, \tfrac{1}{2}\lambda.$$

At a point any distance $x$ from N the path retardation is $\delta$ where

$$\frac{\delta}{\frac{1}{2}\lambda} = \frac{x}{a}, \quad \therefore \delta = \tfrac{1}{2}\lambda\frac{x}{a},$$

and we have the relation

$$(e - e')(\mu_e - \mu_o) = \frac{x}{a} \cdot \frac{\lambda}{2}.$$

At points where $x/a$ is not a whole number, the phase difference will be other than a multiple of $\pi$, and the transmitted light will in general be elliptically[1] polarised.

[1] At points half-way between those where the plane polarisation occurs, the phase difference will be 90° or an odd multiple of 90°, hence if the incident light be polarised at an angle of 45° to the axis of the quartz, the component vibrations will

Hence if the compensator be viewed through a Nicol's prism, turned so as to extinguish the plane-polarised light of the central line N, it will also extinguish the light from a system of parallel lines on either side of N, at a distance $2a$ from each other. Between these dark lines there will be a certain amount of illumination. At points half-way between them—that is, on a system of lines distant from N by odd multiples of $a$—the light is also plane-polarised, but at an angle $2\alpha$ to that at the central line. The Nicol will not extinguish this plane-polarised light but will transmit more or less of it according to the value of $\alpha$, and (if $\alpha = 45°$) the light from these lines will be completely transmitted by the Nicol, so that they will be very bright and the fringes will then be best marked, for between the dark lines corresponding to the $x = 2na$ we shall have the spaces the brightest possible. Hence if $\alpha = 45°$—that is, if the incident light is polarised at an angle of 45° to the principal plane of the face—the fringes will be most distinct; the dark bands being given by

$$x = 2na,$$

and the brightest lines by

$$x = (2n + 1)a.$$

Thus in general we have two systems of lines along which the light is plane-polarised, one corresponding to retardations equal to even multiples of $\frac{1}{2}\lambda$ and the other to odd multiples, and the planes of polarisation of the two systems are inclined at an angle $2\alpha$. If the principal plane of the analyser be perpendicular to the primitive plane of polarisation—that is, if the Nicols be crossed—a system of dark lines corresponding to the former will appear in the field, and if the analyser be rotated through an angle $2\alpha$ a system of dark lines corresponding to the latter will be presented, while for intermediate positions of the analyser there will be no perfectly black bands, but merely lines of maximum and minimum intensity.

The distance $a$ may be measured by fixing a very fine cross wire, furnished with a micrometer screw, so as to coincide with the central dark band. By turning the screw the wire is displaced to the next dark line, and the distance $2a$ through which it has been displaced may be found; or the wire may be displaced to the $n$th dark line from the centre, the distance $2na$ read off, and $a$ calculated. The retard-

be of equal amplitude, and when they differ in phase by 90° the light will be circularly polarised. There will therefore in this case be a system of lines along the face of the plate at which the transmitted light is circularly polarised, and these lie half-way between the lines of plane polarisation. They may be detected by introducing a quarter-wave plate or Fresnel's rhomb, which will reduce the circularly polarised light to plane-polarised light; this may be quenched by a Nicol's prism, and a corresponding system of dark lines will be presented in the field.

ation at any point depends on the wave length, and consequently if white light be used the bands will be rainbow-coloured and the central line alone will be quite black.

*Cor.*—If the angle ABD of one of the prisms be $i$ we have

$$(\mu_e - \mu_o) \tan i = \lambda/4a,$$

for

$$\tan i = \frac{e - e'}{2x}, \quad \text{and } (e - e')(\mu_e - \mu_o) = \frac{x}{a}\frac{\lambda}{2},$$

which by multiplication give the above result. Knowing $a$ and $i$, this gives a method of determining either $\lambda$ or $\mu_e - \mu_o$, when one of them is known otherwise.

**244. Elliptically Polarised Light.**—M. Jamin applied Babinet's compensator, in a modified form, to the determination of the constants which characterise an elliptically polarised ray. In the foregoing we have supposed the cross wire movable while the compensator remains fixed, but in M. Jamin's apparatus the cross wire remains fixed and one of the wedges forming the compensator is moved parallel to itself by means of a micrometer screw, the other wedge remaining fixed. The effect of this is to diminish or increase the difference of thickness $e - e'$ under the wire according to the direction of motion. The motion of the wedge consequently displaces the whole system of fringes across the field, and any particular band may be brought under the cross wire. It is clear that the distance through which the wedge must be displaced in order to displace the system through the width of a band is twice as great as the corresponding displacement $2a$ of the fibre in the first form of the apparatus. For here the thickness of one wedge under the wire remains constant while the other varies, but in the case of a movable cross wire the thickness of one increases and that of the other simultaneously diminishes by the same amount, so that to produce the same difference of thickness under the fibre it is only necessary to move it through half the amount. Hence if $2b$ be the distance through which it is necessary to displace the wedge in order to displace the fringes under the cross wire by the width of a band, we have for the retardation at a distance $x$ from the central band

$$\frac{\delta}{\lambda} = \frac{x}{2b}, \quad \text{or } \delta = \frac{x}{b}\frac{\lambda}{2},$$

and therefore

$$b = 2a.$$

The instrument may now be applied to the determination of the characteristics of elliptically polarised light.

*Determination of the Phase Difference.*—If the incident light be elliptically polarised we may resolve the vibration at each point into two components parallel to the axes of the quartz wedges. These components will differ in phase and amplitude, and are of the general form

$$x = \mathrm{A}\cos(\omega t + \alpha), \quad y = \mathrm{B}\cos(\omega t + \beta).$$

The effect of transmission through the plate is to change the phase difference $\alpha - \beta$ by an amount

$$\delta = \frac{2\pi}{\lambda}(e - e')(\mu_e - \mu_o),$$

which will vary from point to point. It is clear then that we will have a system of parallel lines for which the whole phase difference $\alpha - \beta + \delta$ will be equal to multiples of $\pi$, and the transmitted light along these lines will be plane-polarised, consequently when the compensator is viewed through a Nicol, properly placed, a system of bright and dark bands will be exhibited in the field, and under the central band the total phase difference will be zero.

Now let us suppose that the apparatus is so adjusted that with plane-polarised light the central band is under the fibre. Then, when the incident light is elliptically polarised, the central band will not be under the fibre but will be displaced across the field. At this band the original phase difference existing between the component vibrations has been neutralised by that introduced by the plate. This phase difference is consequently determined by turning the micrometer screw till the central dark band is brought under the wire. If the displacement be $x$ we have therefore

$$\frac{\alpha - \beta}{\pi} = \frac{x}{b}, \quad \text{or } \alpha - \beta = \pi\frac{x}{b}.$$

*Position of the Axes.*—The direction of the axes of the elliptic vibration may also be determined by means of the compensator. We know that the phase difference of the component vibrations taken along the axes is 90° (see Art. 47). Hence set the compensator so that the cross wire is over the central line N—that is, over the central dark band when the incident light is plane-polarised—and turn the screw by an amount $\frac{1}{2}b$, so that the line under the fibre will correspond to a retardation of $\frac{1}{4}\lambda$ or a phase difference of 90°. Now allow the elliptically polarised light to fall on the compensator, and if the central black band be not under the fibre turn the compensator round the direction of the incident light, and so cause the fringes to move across the field, till the central band is under the fibre. In this position the axes of the elliptic vibration are parallel to those of the quartz wedges.

*Ratio of the Axes.*—When two rectangular vibrations compound into another rectilinear vibration, the tangent of the angle which the direction of the resultant makes with one of the components is measured by the ratio of the amplitudes of the components (Art. 47, Cor. 1). Now an elliptic vibration is divided into two components parallel to the axes of the compensator, and at the black bands their phase difference is a multiple of $\pi$ so that they compound into a rectilinear vibration, which, since it is extinguished, must be perpendicular to the principal plane of the analyser; consequently the principal plane of the analyser makes with one of the principal sections of the compensator an angle the tangent of which is equal to the ratio of the amplitudes of the components of the elliptic vibration. Hence if the compensator be set so that its axes are parallel to the axes of the elliptic vibration, then the tangent of the angle between one of its principal planes and the principal plane of the analyser will be equal to the ratio of the axes of the elliptic vibration.

**245. Application of the Quarter-wave Plate.**—A simple but much less precise method of analysing elliptically polarised light is to reduce it to plane-polarised light by transmission through a quarter-wave plate. The property of such a plate is to introduce a phase difference of a quarter period (90°) between the component vibrations which are transmitted through it. Hence if the light be received normally on the plate, placed so that the directions of vibration in it are parallel to the axes of the ellipse, the component vibrations in the transmitted light will differ in phase by 180°, and it will therefore be plane-polarised, and capable of being extinguished by a Nicol's prism. Hence the plate and Nicol are adjusted by trial till the light is extinguished, and in this position we know that the principal sections of the plate are parallel to the axes of the elliptic vibration, and further, the principal plane of the analyser being perpendicular to the direction of vibration of the plane-polarised light emerging from the plate will make an angle with one of the principal sections of the plate, the tangent of which is equal to the ratio of the axes of the elliptic vibration.

Since the retardation in passing through a plate depends on the wave length, it follows that a quarter-wave plate will only act as such for some particular wave length. Fresnel's rhomb, on the other hand, is almost quite free from this objection, and introduces a phase difference of 90° very approximately for all values of $\lambda$. It is consequently better adapted for work with white light.

**246. The Strain Compensator.**—A slip of glass when stretched (or compressed) acquires the doubly refracting properties of a uniaxal crystal of which the optic axis is parallel to the direction of stretching. When interposed in the path of a pencil of plane-polarised light the glass plate will therefore in general convert it into elliptically polarised light, and conversely it may, with proper stretching, reduce to zero, or $\pi$, the phase difference already existing in an elliptic vibration, and so convert it into plane-polarised light. In this application it was employed by Dr. Kerr[1] in his celebrated experiments on the reflection of plane-polarised light from the pole of a magnet (Art. 263).

### *Methods of producing and identifying Polarised Light*

**247. General Method of Detection.**—It may be useful here to recapitulate the methods by which the various kinds of polarised light may be obtained, and to point out the special characteristic qualities by which they may be distinguished from each other, and from common light.

Polarisation of any kind may be detected by means of the colours produced by transmission through a thin crystalline plate, and a knowledge of the character of the fringes afforded by the different kinds of polarised light would enable us by this method alone, not only to detect the presence of the polarisation, but also to determine its class. Each kind may, however, be studied separately as follows.

**248. Plane Polarisation.**—Plane-polarised light was first obtained by double refraction and afterwards by reflection from glass. In all cases of double refraction the refracted pencils are both plane-polarised, and their planes of polarisation are at right angles. An obvious method of obtaining a beam of plane-polarised light is to transmit ordinary light through a doubly refracting crystal and to intercept one of the refracted pencils. This may be done in three ways. If the crystal be large, so as to afford considerable separation of the refracted rays, one of them may be stopped by an opaque diaphragm, while the other is allowed to pass for use or examination. If the separation of the rays be small, so as not to permit of the use of a diaphragm, one of the pencils may be absorbed and the other transmitted. This happens in some natural crystals, such as tourmaline, which possesses a high coefficient of absorption for the ordinary ray, but freely transmits the extraordinary. Finally we possess the method commonly used in practice, viz. that of intercepting one of the beams by total reflection while the other is transmitted, and on this principle such instruments as Nicol's and Foucault's prisms are constructed.

Methods of producing.

[1] Kerr, *Phil. Mag.* November 1875.

The second method of obtaining plane-polarised light, viz. by reflection at the polarising angle from some transparent substance, such as glass, is very convenient, requiring no special apparatus, but it is less perfect than the method of double refraction, for it is found by experiment that very few substances completely plane-polarise light by reflection, and that these have a refractive index of about 1·46.

It is to be remembered that any instrument which acts the part of a polariser may also be used as an analyser to examine the properties of a polarised beam. Examined by any analyser, such as Nicol's prism, the characteristic of plane-polarised light is that it may be completely extinguished by the analyser placed in a certain azimuth, viz. such that the principal plane of the analyser is parallel to the plane of polarisation of the incident light. Examined by transmission through a doubly refracting rhomb, we obtain in general two refracted pencils of different intensities, and in two positions of the rhomb one of the pencils vanishes so that the incident light is transmitted in a single beam. We consequently possess the means of detecting it, and distinguishing it from all other kinds of polarised light, and from ordinary light. Methods of detecting.

**249. Circular Polarisation.**—A circular vibration arises when two rectangular plane-polarised (that is, rectilinear) vibrations of the same period are compounded, which differ in phase by 90° and are of equal amplitude. The equations of such a vibration may be written in the form

$$x=a\cos\omega t,\quad y=a\sin\omega t \qquad \text{(right-handed)},$$

in which case the rotation is from right to left (see Fig. 21), and the vibration is said to be right-handed. Or the equations may be

$$x=a\cos\omega t,\quad y=-a\sin\omega t \qquad \text{(left-handed)},$$

in which the rotation is from left to right, and the vibration is said to be left-handed.

Circular polarisation is therefore produced when a plane-polarised ray is divided into two others, vibrating at right angles, equal in amplitude and differing in phase by 90°. This is obtained by transmitting plane-polarised light through a quarter-wave plate, the axes of the plate being inclined at 45° to the plane of polarisation of the incident light. The same effect is produced with much less colouring by transmission through Fresnel's rhomb, the plane of reflection being inclined at 45° to the plane of polarisation. Produced.

When circularly polarised light is examined through a Nicol the illumination of the field is independent of the orientation of the Nicol.

2 E

In this respect it resembles ordinary light, but the two may be easily distinguished. For if in the path of a circularly polarised beam we interpose a quarter-wave plate or a Fresnel's rhomb, a further quarter-period difference of phase will be introduced between the component vibrations, and this will either add to or destroy the original phase difference. We have consequently emerging from the quarter-wave plate two rectangular vibrations, of which the phase difference is either zero or $\pi$, and these compound into a rectilinear vibration, so that the light is plane-polarised. This may be completely extinguished by a Nicol's prism, and we have the means of deciding whether the original light was circularly polarised, or merely a beam of ordinary light; for the latter on transmission through the quarter-wave plate will still retain the character of ordinary light and will never be quenched by the Nicol.

Detected.

The quarter-wave plate and the Fresnel's rhomb consequently not only act as producers of circular polarisation, but also assist in detecting it where it already exists.

**250. Elliptic Polarisation.**—The rectangular components of an elliptic vibration differ both in phase and amplitude. The equations of the vibration may be written in the general form

$$x = a\cos(\omega t + \alpha), \quad y = b\cos(\omega t + \beta) \qquad \text{(right-handed)},$$

in which case the vibration is said to be right-handed (see Fig. 16), or in the form

$$x = a\cos(\omega t + \alpha), \quad y = -b\cos(\omega t + \beta) \qquad \text{(left-handed)},$$

when the vibration is said to be left-handed. An elliptically polarised ray may therefore be regarded as the resultant of two plane-polarised rays vibrating at right angles, and differing both in phase and amplitude. It consequently includes under it, as particular cases, both plane and circular polarisation.

Produced.

It is clear, then, that any process which produces two plane-polarised beams, polarised at right angles, and differing in phase and amplitude, will in general afford elliptically polarised light. This is effected by transmitting plane-polarised light through a crystalline plate so thin as not to cause visible separation of the two refracted beams, and in general by reflection from metallic or crystalline surfaces.

Detected.

Elliptically polarised light when examined through a Nicol will give an illumination which will vary as the Nicol is rotated, being greatest when the principal plane of the analyser is parallel to one axis of the ellipse and least when parallel to the other. It is thus distinguished from ordinary light, but the same property is possessed by a mixture of ordinary and plane-polarised light. It therefore remains to distinguish between elliptic and partial polarisation. This is effected by the help of a quarter-wave plate or a Fresnel's rhomb properly placed. Thus if the light be transmitted through a quarter-wave plate, and if the axes of the plate be placed parallel to those of the elliptic vibration, the phase difference of the components will be reduced to zero, or increased to $\pi$, and the transmitted light will be plane-polarised. This may be extinguished by a Nicol, and the elliptic polarisation is completely distinguished. The further study of the elliptic vibration is effected by means of Babinet's compensator.

**251. Partial Polarisation.**—The term *partially polarised light* is generally applied to a mixture of ordinary and plane-polarised light. When examined through a Nicol the illumination of the field will vary as the Nicol is rotated, and this arises from the suppression of the plane-polarised component of the light. At first sight it might therefore be suspected that such a beam possessed elliptic polarisation, but as already indicated, a quarter-wave plate assists us to distinguish between the two.

**252. Natural Light.**—It now becomes a matter of interest and importance to inquire into the nature of the vibration in common or unpolarised light. The characteristic of common light is that, on transmission through a doubly refracting rhomb, it affords two pencils of light which do not change in intensity as the rhomb is rotated round the direction of the incident pencil. Circularly polarised light also possesses the same quality, but is, in addition, capable of affording coloured fringes when transmitted through a thin crystalline plate and subsequently analysed. Plane and elliptically polarised light also afford such interference fringes, but in the case of common light they are wholly absent.

When transmitted through a thin crystalline plate ordinary light gives two rays polarised at right angles, and each of these gives an image in the analyser; but the two images are complementary in intensity and superposed. The result is that when the analyser is rotated the illumination remains uniform and equal to half that of the incident light. Hence if the vibration of common light be elliptic, circular, or rectilinear, the nature of the vibration cannot remain permanent, for the light would then possess the qualities of polarised light and exhibit fringes when transmitted through a thin crystalline plate.

This characteristic quality of common light has led to the supposition that in it the vibration frequently changes its character, or that it consists of successive series of elliptic vibrations (including circular and plane as particular forms), all the vibrations of any series being similar to each other, but those of one series having no necessary similarity in direction or form to those of any other. We know that many thousand interference bands may be presented when common light is reflected from Fresnel's mirrors (Art. 100), and that a limit to the observable number of these bands is imposed by the limited resolving power of our spectroscopes. Now interference can only be expected between two sources when they emit similar vibrations. Consequently if common light affords many thousand fringes the vibrations emitted by the source must remain similar for as many periods at least.

Hence if we suppose the vibration to change suddenly at frequent intervals, the vibrations of any series between two consecutive changes must be similar to each other, and there must be many thousand vibrations in each series, say a quarter of a million. Now the orange light of the spectrum executes about 500 billion—$500 \times 10^{12}$—vibrations per second, consequently the form of the vibration might change 500 times per second and still have a billion similar vibrations in each set. This would permit of the exhibition of millions of interference bands (a number quite beyond our power of observation), and yet would sufficiently account for the absence of fringes by transmission through a thin plate. For if we have a number of similar vibrations in any direction, these will give all the phenomena of interference rings and colours as long as the vibrations remain similar, and if the vibration changes to the perpendicular direction we will again have coloured fringes, but this time of the complementary character. Hence if the direction or nature of the vibration changes many times per second the several systems will be produced in such rapid succession that only their combined effect will be perceived, and the field will appear uniformly illuminated.

**253. Imitation of Natural Light by the Rotation of Polarised Light.**—If ordinary light be transmitted through a Nicol's prism the emergent light will be plane-polarised, and the vibration will take place in a certain definite direction determined by the position of the principal plane of the prism. If the Nicol be now rotated round an axis parallel to the direction of the incident light, the plane of polarisation will rotate at the same rate, and the transmitted light when doubly refracted will possess, like ordinary light, the property of giving two beams of constant intensity, and uniform illumination will be produced instead of colours and rings, when transmitted through a thin crystalline plate. It would be possible, however, to distinguish between a rapidly rotating plane-polarised pencil and ordinary light. For if the rotating beam be examined through a Nicol which is rotated round an axis parallel to the direction of the incident light at a rate equal to that of the polarised beam, then if the principal plane of the analyser be parallel to the plane of polarisation of the ray in one position it will remain so permanently, and the light will be completely quenched.

If a thin plate of mica be attached to the polariser the light emerging from the mica plate will in general be elliptically polarised, and the combined rotation of the Nicol and plate will give an elliptically polarised beam, of which the vibrations are similar but rotating rapidly, so that the axes of the ellipses are directed in all azimuths. Such a beam was found to produce[1] the same polarisation phenomena (such as coloured rings, etc.) as circularly polarised light, but if the Nicol be kept fixed and the mica plate rotated, so that the sense of the elliptic vibration is reversed every half revolution, the transmitted pencil behaved like common light.

It has been shown (Art. 48, Ex. 2) that a plane-polarised ray may be considered as the resultant of two opposite circularly polarised rays of the same period. If the circular vibrations be of different periods the direction of the resultant vibration will rotate uniformly, and therefore a rotating plane-polarised ray, such as we have considered above, may be regarded as the resultant of two opposite circularly polarised rays of different periods. This difference in period means a difference in wave length and refrangibility, and Airy[2] has consequently remarked that a gradual change in the vibration, such as the uniform rotation of its direction, is not an admissible representation of the nature of common light, for we may regard a uniformly rotating vibration as the equivalent of two others of different periods, and it should give rise to two polarised rays of different colours. The velocity of rotation required to produce any such separation would, however, be far in excess of anything we could possibly attain.

Rotating ray.

**254. Haidinger's Brushes.**—It was discovered by Haidinger that plane-polarised light may be detected by the naked eye, and its plane of polarisation may also be ascertained. The phenomenon observed is the appearance of a pale yellow patch or brush bounded on either side by curved arcs in form like the branches of a hyperbola. The axis of this brush lies in the plane of polarisation, and on each side of its neck there is a violet or bluish patch.

These brushes are supposed to arise from the polarising structure of the eye itself, and by some persons they are observed even when the light is only partially polarised. Most persons, however, can see them only

[1] Dove, *Pogg. Ann.* vol. lxxi. p. 97. See Verdet, *Œuvres*, tom. vi. p. 88.
[2] Airy, *Undulatory Theory of Optics*, Art. 183.

when the light is completely polarised, and even then with difficulty. They are best observed by looking at a bright cloud through a Nicol's prism, the Nicol being revolved round its axis so as to alter the plane of polarisation. They gradually disappear when the eye is directed towards them in the same position, being visible for only a couple of seconds after the polarised light is first received by the eye; but when the position of the plane of polarisation is changed by rotating the Nicol the brushes continue in the field and are seen to revolve with the Nicol. They may be seen by looking for a few moments at one of the images afforded by a rhomb of Iceland spar and then at the other, and so on alternately, or by looking at the sky, alternately towards the horizon and the zenith, in a plane at right angles to the line joining the observer to the sun when the sun is near the horizon.

M. Jamin attributes this phenomenon to the refracting coats of the eye. Thus if plane-polarised light be transmitted through a pile of plates the intensity of the transmitted beam will depend on the relation of the plane of incidence to the plane of polarisation. Now the crystalline lens of the eye consists of curved layers of unequal refracting powers, and, on account of the curvature, the relation of the plane of polarisation of a given pencil of polarised light to the plane of incidence at each point will vary from point to point, with the result that the intensity of the light transmitted will vary from point to point on the retina, and the general appearance of the brushes would be accounted for. Helmholtz,[1] however, took objection to this explanation, and found that in working with homogeneous light[2] the blue was the only colour with which the brushes are visible, and that the extent of the brushes is limited to the yellow spots of the eye. To account for the brushes it therefore suffices to suppose that the yellow spots are doubly refracting to a slight extent, and that they absorb the extraordinary ray more strongly than the ordinary in the case of blue light.

**255. The Polarisation of Skylight.**—It has been already noticed that, in general, polarisation is produced to some extent by reflection and refraction, and it is consequently not surprising that the light of the rainbow, and even the light of the open sky, should be more or

[1] *Physiological Optics.*

[2] The office of the different colours in the production of these brushes was first investigated by Sir G. G. Stokes, who found that in the red and yellow light of the spectrum no trace of them could be observed. The brushes began to be visible in the green, were more distinct on passing into the blue, were particularly strong about the line F, could be traced as far as the line G, and when they were no longer visible the cause appeared to be merely the feebleness of the light, not the incapacity of the greater part of the violet to produce them (Stokes, *Brit. Assoc. Report*, 1850; *Collected Papers*, vol. ii. p. 362).

less polarised. The polarisation of the blue light of the sky was noticed by Arago as early as 1811. When the open sky is viewed through a Nicol's prism variations of brightness are observed as the Nicol is rotated on its axis. This change of brightness is most noticeable when the sky is viewed in a direction at right angles to the rays of the sun, and observation proves that the light of the sky is polarised, and that the direction of most perfect polarisation is at right angles to the direction of propagation of the solar rays. This is sometimes expressed by saying that the place of maximum polarisation in the sky occurs at an angular distance of 90° from the sun.[1]

The blue light of the sky has been imitated, and artificial skies produced in a most elegant manner by the late Professor Tyndall. If the peculiarities of skylight are due to the scattering action of very small particles suspended in the air, the object of a test experiment must be to procure a region of space in which very fine particles are known to be suspended and to allow a beam of light to play upon it. This was secured by passing a pencil of light through a tube containing a small quantity of vapour of iodide of allyl.[2] By the action of the light the vapour is gradually decomposed, and a very fine cloud forms within the tube. At first the precipitated particles are exceedingly small, and the colour observed is a delicate blue, but as the experiment progresses the particles gradually increase in size, and the blue gradually brightens, "still maintaining its blueness, until at length a whitish tinge mingles with the pure azure, announcing that the particles are now no longer of that infinitesimal size which scatters only the shortest waves."[3] Further, when the incipient cloud was viewed transversally through a Nicol's prism, the scattered light was found to be completely polarised in the plane passing through the axis of the tube and the eye of the observer—that is, the plane containing the incident and scattered ray—just as the blue light of the sky is polarised in a plane passing through the solar rays and the eye of the observer.

That the preponderance of blue in the light of the sky is due to the scattering action of very small particles suspended in the air seems to be placed beyond doubt by these experiments. Now we have seen that this preponderance becomes at once intelligible when we consider

[1] In the vertical plane through the sun neutral points have been observed by Arago, Babinet (*Comptes Rendus*, tom. xi. p. 618, 1840), and Brewster (*Brit. Assoc. Report*, 1842, part ii. p. 13). At these points there is no polarisation, and in the intervals between them the light is polarised in rectangular planes.

[2] Several other substances will also produce the desired effect, as nitrite of amyl, bisulphide of carbon, benzole, benzoic ether, etc.

[3] Tyndall, *Heat a Mode of Motion*, p. 491.

the action of particles of linear dimensions small compared with the wave length of light (Art. 162), and Lord Rayleigh[1] has further shown that the polarisation of the scattered light may be explained by supposing that the particles *load* the ether so as to virtually increase its inertia, or that property of it which corresponds to inertia (Art. 311).

Loaded ether.

This loading action of the particles might be counterbalanced by supposing suitable forces to act at all points of the ether where the inertia is altered. If these forces were applied the loading would be neutralised, and the waves would pass unbroken just as if the particles were not present, so that there would be no scattering of the light. These applied forces must have the same period and direction as the undisturbed luminous vibrations, and the light actually scattered is the same as would be caused by the action of forces exactly the opposite of the foregoing acting on the medium otherwise free.

Now on account of the smallness of the particles we may take the forces acting throughout the volume of any one of them as a whole. The periodic disturbance arising from the action of this force will be symmetrical round the direction of the force, and the transverse vibrations which it originates will be such that the direction of vibration in any ray lies in the plane containing the ray and the axis of symmetry; in other words, the direction of vibration in the diffracted ray makes the least possible angle with the direction of vibration in the incident ray. Further, the intensity of the scattered light should vanish in the direction of the axis of symmetry, for the transversal to a ray propagated in this direction becomes indeterminate. This being conceded, let us suppose that a solar ray is regarded transversely, so that the line of vision is at right angles to the ray. The vibration in the ray may in this case be divided into two components, one along the line of vision and the other at right angles to it, both of these components being in the wave front. Now the component along the line of vision will give rise to no scattered light in this direction, for this is its axis of symmetry, and consequently the scattered light received by the eye arises entirely from the component at right angles to the line of vision. It is therefore completely polarised, and the plane at right angles to the direction of vibration—that is, the plane of polarisation in Fresnel's theory—contains the line of vision and the incident ray.

When the angle between the line of vision and the incident ray is other than 90° the scattered light is a mixture arising from both the foregoing components, and the polarisation becomes less and less complete as the line of vision approaches the direction of the ray, and in this direction it altogether disappears.

[1] Hon. J. W. Strutt, *Phil. Mag.* vol. xli. p. 107, 1871.

## CHAPTER XVII

### ROTATORY POLARISATION [1]

**256. Rotation of the Plane of Polarisation—Arago's Discovery.**—When plane-polarised light is reflected at the surface of a transparent substance, the planes of polarisation of the reflected and refracted pencils do not in general coincide with that of the incident light (Art. 210). If the plane of polarisation of the reflected light is inclined at an angle $\theta$ to that of the incident, we say that the reflection has rotated the plane of polarisation through an angle $\theta$, and this rotation is said to be right-handed or left-handed according as it is (to an observer receiving the light) from right to left, or the reverse. Looking in the direction in which the light is travelling, a right-handed rotation will consequently be from left to right, like the motion of a right-handed screw, and a left-handed rotation will be from right to left. In this case the rotation is independent of the length of path traversed by the light before or after incidence. It is therefore due to no action of the medium through which the light passes, but occurs at the surface of separation where the reflection takes place.

In 1811 Arago [2] discovered that rotation of the plane of polarisation occurs when plane-polarised light is transmitted through quartz in the direction of its optic axis, and this property is also possessed by many other substances. The rotation is here due to the action of the medium through which the light passes, and its amount is directly proportional to the thickness traversed. When plane-polarised light passes through a uniaxal crystal, such as Iceland spar, in the direction of the axis, the plane of polarisation of the transmitted pencil coincides with that of the incident light, but it is different in the case of rock-crystal. Here the plane of polarisation at emergence is not the same as at incidence, but is inclined to it at an angle depending on Right and left.

[1] [In the author's copy I find the query "Can rotatory polarisation be explained by internal reflection, *i.e.* by Art. 210?"]

[2] *Mém. de la première Classe de l'Institut*, tom. xii. p. 93. *Œuvres Complètes*, tom. x. p. 35, 1812.

the thickness of the quartz plate. Further, some specimens of quartz rotate the plane of polarisation to the right (looking in the direction of propagation of the light) and some to the left. The former are called right-handed or dextrogyrate, and the latter left-handed or levogyrate.

**257. Biot's Laws.**—This remarkable phenomenon was investigated with great care and success by Biot,[1] who deduced the following laws:—

(1) The amount of rotation is proportional to the thickness traversed by the ray.

(2) The rotation effected by two plates is the algebraic sum of the rotations produced by each separately.

(3) The rotation augments with the refrangibility of the light, and is approximately proportional to the inverse square of the wave length.

In support of the third law, M. Broch gives the following numbers for quartz. The first line gives the rotation $\rho$ produced by a quartz plate (one millimetre thick) on the different rays, and the second line gives the product of $\rho$ by the square of the corresponding wave length :—

| Ray | B | C | D | E | F | G |
|---|---|---|---|---|---|---|
| $\rho$ | 15° 30′ | 17° 24′ | 21° 67′ | 27° 46′ | 32° 50′ | 42° 20′ |
| $\rho\lambda^2$ | 7238 | 7429 | 7511 | 7596 | 7622 | 7841 |

For essence of turpentine Wiedemann found

| | | | | | |
|---|---|---|---|---|---|
| $\rho$ | 10° 9′ | 14° 5′ | 18° 7′ | 23° 2′ | 32°·75′ |
| $\rho\lambda^2$ | 4690 | 4871 | 5184 | 5471 | 6044 |

These tables show that the rotation is not strictly in the inverse ratio of the square of the wave length, but that the product $\rho\lambda^2$ increases with the refrangibility of the light. Biot's law is accordingly only an approximation to the truth.

**258. Rotation produced by Liquids and Vapours.**—It was soon discovered that many liquids and solutions possess, like quartz, the power of rotating the plane of polarisation of a ray transmitted through them, but in a much less degree. Thus a plate of quartz 1 mm. thick rotates the plane of polarisation of the red rays about 18°, while an equal thickness of turpentine only produces a rotation of a quarter of a degree.

The rotatory power of liquids and vapours is studied by enclosing them in a tube through which the polarised light is transmitted. It is found that liquids and solutions do not lose their rotatory power when diluted with other liquids not possessing this property, and they retain it even when in the state of vapour. Biot consequently concluded that the power possessed by liquids of rotating the plane of

[1] Biot, *Mém. de la première Classe de l'Institut*, tom. xiii. p. 218, 1813. *Ann. de Chimie et de Physique*, 1815, etc.

polarisation is inherent in the ultimate molecule. In this respect they differ from quartz, which loses its rotatory power when it loses its crystalline arrangement. Thus Sir J Herschel found that quartz held in solution by potash (liquor of flints) was destitute of rotatory power, and the same was observed by Sir D. Brewster with respect to fused quartz.

It is, perhaps, not surprising that crystalline substances should, on account of some special molecular arrangement, possess rotatory power and affect the propagation of light within the mass in a manner depending on the direction of transmission. The loss of this power when the crystalline structure is destroyed, as when quartz is fused, is consequently an event which would be naturally expected, but the possession of it in all directions by fluids and solutions, in which there cannot be any special internal arrangement of the mass of the nature of a crystalline structure, is not a thing which one would have been led to expect beforehand. To Faraday it appeared to be a matter of no ordinary difficulty, and I am not aware that any explanation of it has ever been suggested. It is just possible that the light in traversing a solution in which the molecules are free to move may, on account of some peculiarity of structure, cause the molecules to take up some special arrangement, so that the fluid becomes as it were polarised by the transmission of the light, in a manner somewhat analogous to that in which a fluid dielectric is polarised in a field of electrostatic force.

When two or more liquids are mixed together their combined rotation is always equal to the algebraic sum of the rotations which the constituents would produce separately when taken in thicknesses proportional to the volumes they occupy in the mixture. This property has been applied to the analysis of compounds containing a substance which possesses rotatory power and is combined with others which are neutral. This important application has been found of much industrial value in the analysis of saccharine solutions, and for this purpose instruments termed *saccharimeters* have been invented (Art. 273, etc.).

Calling right-handed rotations positive and left-handed negative, the rotations produced by a thickness of one decimetre of each of the following substances, in the case of red light, are [1]

| | | | |
|---|---|---|---|
| Essence of Turpentine | $-29°{\cdot}6$ | Essence of Seville Orange | $+78°{\cdot}94$ |
| ,, Citron | $+55°{\cdot}3$ | ,, Mint | $-16°{\cdot}14$ |
| ,, Aniseed | $-\ 0°{\cdot}7$ | ,, Sassafras | $+\ 3°{\cdot}53$ |
| ,, Lavender | $+\ 2°{\cdot}02$ | ,, Rosemary | $+\ 3°{\cdot}29$ |
| ,, Fennel | $+13°{\cdot}16$ | ,, Carraway | $+65°{\cdot}79$ |

[1] See Verdet, *Œuvres*, tom. vi. p. 270.

| | | | |
|---|---|---|---|
| Solution of Sugar (cane), | 50 % in water | +33°·64 |
| ,, Quinine | 6 % in alcohol | −30° |
| ,, Strychnine | 1¼% in alcohol | − 6°·6 |
| ,, Brucine | 5 % in alcohol | −12° to −16° |
| ,, Morphine | 5 % in soda | −11°·5 |

For yellow light (D) and one millimetre thickness we have—

| | | | |
|---|---|---|---|
| Quartz | 21°·67 | Hyposulphate of Potash | 8° 39′ |
| Cinnabar | 32°·5 | ,, Lead | 5° 53′ |
| Chlorate of Soda | 3°·67 | ,, Strontium | 1° 64′ |
| Bromate ,, | 2°·80 | ,, Calcium | 2° 9′ |

**259. Molecular Rotatory Power.**—When a substance possessing rotatory power is held in solution by a liquid which is inactive, it is found that rotation of the plane of polarisation is proportional to the quantity of the active substance traversed by the ray—that is, the rotation is proportional to the number of molecules of the active substance in the path of the ray. This has given rise to the opinion that the rotatory power is inherent in the ultimate molecule, and hence the term molecular rotatory power.

If unit weight of the solution contains a weight $w$ of an active substance, and a weight $1-w$ of the inactive solvent, and if the specific gravity of the solution be $\sigma$, the volume of unit weight will be $v=1/\sigma$, and the density of the active substance in the solution will be

$$x=\frac{w}{v}=w\sigma.$$

The rotation produced by a thickness $e$ will therefore be

$$\rho=\mathrm{M}ew\sigma,$$

where M is a constant which measures the rotation produced by unit thickness of a solution containing the active substance in unit density. The constant M is called the *molecular rotatory power of the substance.*

If two active substances be mixed, their weights being in the ratio $w:(1-w)$, the rotation produced by the second will be

$$\rho'=\mathrm{M}'e(1-w)\sigma,$$

where M′ is its molecular constant. Hence their combined rotation will be

$$\rho\pm\rho'=e\sigma\{\mathrm{M}w\pm\mathrm{M}'(1-w)\}.$$

The quantity M is found to be not absolutely constant. In general it augments a little as the proportion of the solvent is increased, and in the case of the organic alkaloids and their salts marked deviations are exhibited.

*Temperature Effect.*—The effect of a change of temperature is generally to change the molecular rotatory power. In some substances, such as essence of oranges and essence of turpentine, the rotatory power is diminished by a rise of temperature, while in others, such as quartz and chlorate of sodium, it increases with the temperature.

M. Gernez[1] expresses the molecular rotatory power in the form

$$M = a - bt - ct^2,$$

where $a$, $b$, $c$ are constants, which diminish rapidly. Thus for the essence of oranges and essence of turpentine M. Gernez finds for the D ray

Essence of orange $M = 115{\cdot}91 - 0{\cdot}1237t - 0{\cdot}000016t^2$,
,, turpentine $M = 36{\cdot}61 - 0{\cdot}004437t$.

In the case of quartz the molecular rotatory power may be represented between $-20°$ and $100°$ by the formula

$$\text{Quartz } M = M_0(1 + 0{\cdot}000146324t + 0{\cdot}0000000329t^2),$$

where $M_0 = 21{\cdot}658$.

**260. Dispersion of the Planes of Polarisation—Tint of Passage.** —Since by Biot's third law the rotation is different for the different wave lengths, it follows that when white light is used the planes of polarisation of the various constituent colours in the transmitted pencil will have suffered different amounts of rotation, and consequently, if this light be examined by means of a Nicol's prism, only one of its constituents will be quenched, viz. that colour for which the plane of polarisation is parallel to the principal plane of the Nicol. The other colours will be more or less transmitted according to the position of their planes of polarisation. The field will consequently appear coloured and the colour will vary as the Nicol is rotated. If the thickness of the quartz plate does not exceed 5 mm., then for a certain position of the Nicol the field possesses a grayish-violet colour, called the *tint of passage* (named by Biot the *teinte sensible*). If the Nicol be slightly turned from this position the field becomes red, and for a rotation in the opposite direction it appears blue. On passing through this position the transition from red to blue is very rapid. It is accordingly a position attainable with considerable accuracy.

When the tint of passage has been obtained for a quartz plate, the principal plane of the Nicol is parallel to the plane of polarisation of the greenish-yellow rays, and these rays are consequently quenched by the Nicol and absent from the emergent light. This point may be

[1] M. Gernez, *Ann. de l'École Normale*, tom. i. p. 1.

tested by submitting the transmitted light to spectroscopic analysis. The spectrum will be found to be crossed by a dark band corresponding to the missing rays.

When the analyser is placed so as to produce the tint of passage, the illumination of the field is a minimum, for the rays corresponding to the brightest part of the spectrum are absent and the illumination of the field is produced by the less intense parts, viz. the red and violet ends of the spectrum. A general expression for the intensity corresponding to any position of the analyser is easily obtained. Thus if the intensities of the various colours in the incident light be $I_r$, $I_o$, $I_y$, etc., for the red, orange, yellow, etc., and if the primitive plane of polarisation of the light make an angle $a$ with the principal plane of the Nicol, then after passing through the quartz they will make angles $a + a_r$, $a + a_o$, etc., where $a_r$, $a_o$, etc., are the angles through which the planes of polarisation of the corresponding colours (red, orange, etc.) are rotated by the quartz. Hence the intensity of the red in the field of the analyser will be $I_r \sin^2(a + a_r)$ with corresponding expressions (see Art. 174) for the other colours. The complete expression for the illumination will consequently be

$$I = I_r \sin^2(a + a_r) + I_o \sin^2(a + a_o) + \quad . \quad . \quad . \quad + I_v \sin^2(a + a_v).$$

**261. Relation of Rotatory Power to Crystalline Form.**—The curious fact that some specimens of quartz rotate the plane of polarisation to the right, while others rotate it to the left, was found by Sir J. Herschel to be intimately connected with a difference of crystalline form. The ordinary form of a quartz crystal is a six-sided prism topped by a six-sided pyramid. The solid angles at the junction of the prism and pyramid are often absent, and replaced by facets, or small secondary planes, which are obliquely inclined to the other faces of the crystal. Crystals presenting this anomaly are named *plagihedral*. In the same crystal the planes all lean in the same direction, and when that direction is to the right, the apex of the pyramid being uppermost, the crystal is right-handed, and when it is to the left the crystal is left-handed.[1]

[1] Sir David Brewster subsequently discovered that the *amethyst*, or violet quartz, is made up of *alternate layers* of right-handed and left-handed quartz. This remarkable structure may be traced in the fracture of the mineral, for the edges of the layers *crop out*, and give to the fracture the undulating appearance which is peculiar to the mineral. But the structure in question is displayed in the most beautiful manner when we expose a plate of this substance to polarised light. The colours exhibited in polarised light likewise reveal the existence of crystals of quartz penetrating others in various directions, when no striæ, or other external appearances, indicate their presence (Lloyd, *Wave Theory of Light*, p. 239).

**262. Rotation produced in a Magnetic Field—The Faraday Effect.**—In 1845 Faraday[1] made the remarkable discovery that isotropic substances, more especially those possessing a high refractive power, such as heavy glass, may also acquire the property of rotating the plane of polarisation when under the influence of powerful magnetic force. In Faraday's experiment the soft iron pole-pieces of an electro-magnet were each pierced with a cylindrical hole, and a pencil of plane-polarised light was transmitted through them, the pole-pieces being placed so that the axes of the holes were in the same straight line. The direction of this line and the direction of the transmitted pencil of light were consequently parallel to the lines of magnetic force when the magnet was excited. Between the poles, and in the path of the light, a block of dense glass (silicated borate of lead) was placed, and the light was received by an analyser, adjusted so that the field was dark.[2] On allowing a current to pass in the coils of the magnet, a powerful magnetic field was produced and the light reappeared in the analyser, but it could be again extinguished by properly turning the analyser. This showed that the pencil emerging from the glass was still plane-polarised, but that the plane of polarisation was rotated by the influence of the magnetic field. When the current was reversed the direction of magnetisation was changed, and it was found that the direction of rotation of the plane of polarisation was also reversed.

The direction in which the rotation of the plane of polarisation takes place is determined by a rule similar to Ampère's law connecting the direction of the lines of magnetic force encircling an electric current with the direction of the current. Thus, as Faraday states it, if a watch be placed between the poles of the magnet, so that its face is towards the north pole and its back towards the south, then the direction of Rule. motion of the hands of the watch is the direction of rotation of the plane of polarisation; in other words, to a person looking along a line of magnetic force—that is, from north to south, or from places of higher to places of lower magnetic potential—the rotation is from left to right. Or, again, if we regard a line of magnetic force as due to an

[1] Faraday, *Experimental Researches* (xix$^{th}$ series), vol. iii. p. 1; *Phil. Trans.* 1846, p. 1.

[2] It is difficult to secure blocks of glass free from internal stress, and, as already remarked (Art. 224), glass in this condition plays in some degree the part of a doubly refracting crystal and "depolarises" the light. It is accordingly important that the block of glass used in this experiment should be well annealed or else corrected in some way by a compensating plate. The silicated borate of lead glass softens at a temperature below that of boiling oil and can be well annealed. Flint glass possesses a more feeble rotatory power, and crown glass is more feeble still.

electric current encircling that line, then the direction of rotation is the same as the direction in which the current circulates. A rotation in this direction is consequently referred to as a *positive* rotation. When the light passes through the field in a direction parallel to the lines of magnetic force the effect is greatest, and when the direction of the light is perpendicular to the lines of force there is no rotation of the plane of polarisation.

Faraday's experiments extended over a vast variety of substances which included solids, liquids, and gases, but in the latter he was not able to observe any decided effect. In many of these experiments the electromagnet was dispensed with, and the magnetic field produced within a long helix of wire, carrying an electric current, was employed instead. By this means the light could be transmitted through a great thickness of any transparent substance, placed as a core within the helix, and feebly rotating substances could be examined with more effect. In all cases the magnetic field of the solenoid behaved as that of the electromagnet, except that with the former the light in the field of the analyser was suddenly restored when the current was switched on, whereas with the latter the restoration of the light was more or less gradual.

In the case of substances in which the induced rotatory power was feeble, it was found best to work with the Nicols turned a little from the position of complete extinction, so that with no current there was a little illumination in the field of the analyser. When the current was turned on in one direction this illumination was increased, whereas in the other direction it was diminished, and this change of intensity in opposite directions doubled the effect to be observed.[1]

The property thus induced by the magnetic field is similar to that possessed by saccharine solutions and other rotatory substances, in so far that a rotation of the plane of polarisation is effected, but there is one important difference. In the case of rotatory substances the direction of rotation is the same whatever be the direction of transmission—that is, if a ray be rotated to the right in passing from one end A to the other end B of a crystal of quartz, it will also be rotated to the right in passing from B to A. The result is, that if a ray pass

Distinction.

[1] Faraday noticed that the rotation produced by a substance (water) was apparently unaffected by placing an iron bar within the helix, and that it was the same whether it was placed along the axis or near the side of the spiral. It was also the same whether the substance was enclosed in an iron or a brass tube. When water was placed in an iron tube $\frac{1}{8}$ inch thick, and this tube placed within another of the same thickness but a little wider diameter, the rotation was increased. On placing these two within a third iron tube the effect diminished, but was still very considerable.

from A to B, and be then reflected directly back so as to pass from B to A, the rotations will neutralise each other, and the ray after the double transmission will return polarised in the primitive plane of polarisation. This is easily conceived, for if the observer be supposed looking in the direction AB, and if the incident ray in passing from A to B be rotated to his left (that is, to the right of an observer looking in the direction BA who would receive the emergent ray), then the reflected ray in returning from B to A will be rotated by an equal amount to his right, so that the two will neutralise each other. In the case of the magnetic field, however, the direction of the rotation depends on the direction of transmission, as well as on the nature of the medium. If the rotation is to the right in passing from a place of higher to a place of lower potential, then it will take place to the left when passing from lower to higher. It follows, therefore, that if a ray pass from A to B and be then reflected back so as to pass from B to A, the rotation of the plane of polarisation will be doubled. In the magnetic field the rotation is in the same direction whether we conceive the ray as moving from us or towards us, but with a quartz crystal, if the rotation be to the right for a ray advancing towards us, it will be to the left for one travelling from us.

This difference in character, of the rotation produced by a quartz crystal and that effected by a substance placed in a strong magnetic field, points to a difference in the nature of the causes which primarily induce the rotation. The fact that the rotation is removed by retraversing a piece of quartz, or any natural rotatory substance, indicates something analogous to a twisted structure in the substance, but the doubled effect of the magnetic field seems to indicate that something of the nature of a rotation is going on in the field. This rotation must be an affection of, or at least depends on, the matter occupying the field, and does not necessarily lead to the conclusion that the free ether in a magnetic field is in rotatory motion, for the Faraday effect is in one direction in some substances and in the opposite direction in others. It is consequently induced by the action of the magnetic field on the matter molecules occupying it, the effect of the magnetic action being to originate something analogous to a rotation in the matter molecules, and the direction of rotation being determined by the nature of the matter.

The rotation is found to be proportional to the difference of magnetic potential between the points where the ray enters and leaves the medium. Thus if $V_1$ and $V_2$ be the values of these potentials, we have for the rotation

$$\rho = c(V_1 - V_2),$$

where $c$ is a constant (known as *Verdet's constant*) depending on the nature of the medium, and is generally positive or negative according as the medium is diamagnetic or magnetic. It follows, therefore, that the rotation will be greatest when the ray travels in the direction of the lines of force, and will vanish in directions perpendicular to these lines, while for oblique directions it will vary as the cosine of the angle of inclination.

Since the time of Faraday this rotatory power has been found to be impressed on most diamagnetic substances, and recently it has been observed in gases as well as in solids and liquids, by Becquerel, Kundt, and others. By reflecting the pencil of light backwards and forwards several times through the field, a feeble rotation may be much amplified and rendered sensible, for if the ray traverses the field $n$ times the rotation will be $n$ times that produced by a single passage.

Results. The results of these experiments prove that most isotropic substances (including solids, liquids, and gases), when subject to magnetic force, rotate the plane of polarisation of the transmitted light in the positive direction,—the positive direction of rotation being taken, as already defined, to be the direction of the Ampèrean current producing the field. Further, a positive rotation occurs when light is transmitted through thin films of magnetic substances such as iron, nickel, and cobalt. It is positive also in the case of oxygen (which is magnetic), but in the case of a concentrated solution of ferric chloride the rotation is negative. The negative rotation of other magnetic salts may be recognised by the diminution of the positive rotation of the solvent.[1]

**263. Kerr's Experiments.**—(1) *Electrostatic Effect.*—After discovering the magnetic rotation of the plane of polarisation Faraday sought for a corresponding effect in substances subjected to electrostatic stress, but without success. This effect was first observed by Dr. Kerr[2] of Glasgow, who found that a dielectric under electric stress acquires the double refracting properties of a uniaxal crystal of which the optic axis is in the direction of the lines of electric force, some substances behaving like negative and some like positive crystals. Glass and olive oil and quartz belong to the former class, while bisulphide of carbon, paraffin, oil of turpentine, and resin belong to the latter.

In glass the effect was well marked, and the experiment was conducted by drilling two holes in opposite ends of a block of the material. Metallic terminals were inserted in these holes and con-

[1] A. Kundt, *Phil. Mag.* vol. xviii. p. 327, 1884.

[2] J. Kerr, *Phil. Mag.* vol. l. pp. 337-446, 1875.

nected to the electrodes of an induction coil, which possessed a sparking distance of from 20 to 25 cm., and this distance could be varied at pleasure. The thickness of glass between the terminals was about a quarter of an inch clear, and across this space, at right angles to the lines of electric force, a pencil of plane-polarised light was transmitted. When the induction coil was out of action the transmitted light was plane-polarised and could be extinguished[1] by a Nicol properly adjusted. This position being accurately secured it was found that when the coil was thrown into action, so that the glass was under electric stress, the light reappeared in the field of the analysing Nicol; nor could it be again quenched by any rotation of the Nicol in either direction. The light thus becomes elliptically polarised by the dielectric, and the character of the elliptic vibration may be examined by means of a Babinet's compensator (Art. 243).

The effect is regular and best marked when the light crosses the lines of force at right angles, and the plane of polarisation is inclined to them at 45°; but when the plane of polarisation is either parallel or perpendicular to the lines of force there is no effect. The optical effect does not attain its full intensity at once, but increases gradually from zero to its final state in about thirty seconds after the electric field is excited, and in the same manner the optical effect fades away gradually when the electric stress is removed.

This effect may be completely compensated by a slip of glass compressed, or stretched, in the direction of the lines of force (see Art. 246). From this it is inferred that the effect of the electric action is optically equivalent to compressing the glass in the direction of the electric force, the glass acquiring the properties of a negative uniaxal crystal having its axis parallel to the lines of force. The result is as if the molecules of the glass assumed a regular crystalline arrangement in the electric field, and the effect is as good with the alternating electrifications produced by an induction coil as with a continued electrification of the same strength in one direction produced by a frictional machine; the alternating electrifications produce contrary electric polarisations, but their actions conspire in arranging the molecules.

If $\delta$ be the relative retardation introduced between the components of the elliptic vibration per unit thickness of the dielectric, Kerr[2]

[1] In order that extinction should be produced it was found necessary to introduce a plate of glass in front of the analyser to neutralise the "depolarising" effect of the experimental block containing the terminals. This was called the neutralising block and was adjusted by trial.

[2] Kerr, *Phil. Mag.* vol. ix. p. 157, 1880.

found $\delta$ to be proportional to the square of the electric force F—that is,

$$\delta = kF^2,$$

where $k$ is a constant depending on the nature of the dielectric, and may be either positive or negative. This law contains a result found experimentally—namely, that the optical effect is independent of the direction of the electric force,—that is, it remains the same when the direction of the force is reversed.

(2) *Magnetic Effect.*—Shortly after the discovery of the modifications produced in the optical properties of a dielectric when subject to the action of electric force, Dr. Kerr[1] was led to examine the effects produced when plane-polarised light is reflected from the surface of a magnetised body, such as the polished pole-piece of an electromagnet. In approaching this subject we must remember that when plane-polarised light is reflected from the surface of a polished metal the reflected light is in general elliptically polarised (Art. 219). There are two cases, however, in which the reflected light remains plane-polarised—namely, when the light falling upon the surface is polarised either in or at right angles to the plane of incidence—and in these two cases the reflected light can be extinguished by the analysing Nicol. Let us suppose, therefore, that light polarised in the plane of incidence falls upon the polished pole-piece of an electromagnet (which, for the present, is unmagnetised), and that the reflected pencil is received by an analyser placed in the position of complete extinction. If the magnetising current be now turned on the illumination is observed to reappear in the field of the analyser, and this cannot be extinguished by rotating the analyser in either direction. The reflected light is consequently not plane-polarised. Extinction, however, may be produced by introducing a compensating strip of strained glass into the path of the light, and this shows that the reflected beam is elliptically polarised.

Elliptic polarisation.

This is the new fact discovered by Dr. Kerr, and is described by him as a rotation of the plane of polarisation, thus: "When plane-polarised light is reflected regularly from either pole of an electro-magnet of iron, the plane of polarisation is turned through a sensible angle in a direction contrary to the nominal direction of the magnetising current; so that a true south pole[2] of polished iron, acting as a reflector, turns the plane of polarisation right-handedly."

The arrangement of the apparatus is shown in Fig 196. The source of light was a paraffin flame S, placed at a distance of one foot,

[1] J. Kerr, *Phil. Mag.* vol. iii. p. 321, 1877; and vol. v. p. 161, 1878.

[2] By a *true* south pole is here meant the north-seeking pole of a magnet.

or less, from the polished polar surface. Close to the flame came the polarising Nicol P, through which the light passed in a horizontal direction and fell upon the pole of the magnet, inclined so that the angle of incidence could be varied at pleasure. At a few inches from the pole the reflected light was received by the analyser A, the Nicols being adjusted so that complete extinction was secured when the pole was unmagnetised. In order to obtain any optical effect it was found necessary to secure intense concentration of magnetic force upon the reflecting surface at the point of incidence of the light. For this purpose a block of soft iron, W (one of the several pole-pieces of the magnet), about two inches square and three inches long, was planed off at one end into a blunt wedge with a well-rounded edge. This wedge, called the *submagnet*, was placed over the centre of the reflecting pole, resting on two splinters of hard wood so as to leave a narrow chink through which the light passed, the width of the chink being about $\frac{1}{20}$ of an inch. With this arrangement an intense concentration of magnetic force is produced in that part of the field which is utilised optically, the lines of force being sensibly perpendicular to the reflecting surface.

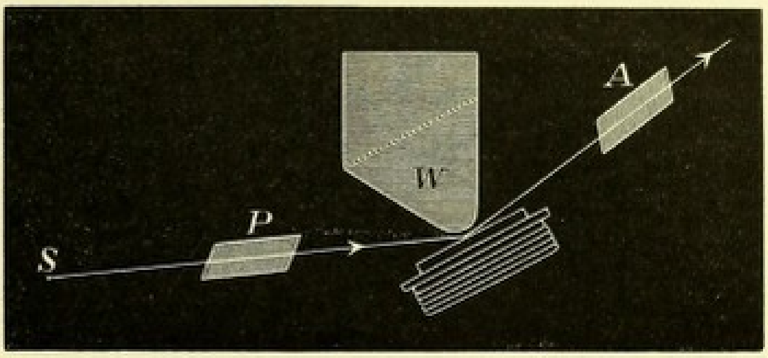


Fig 196.

The submagnet.

In the greater part of his experiments Dr. Kerr (as recommended by Faraday) first adjusted the Nicols to perfect extinction, and then turned the analyser through a very small angle, so that the light was faintly restored in its field. The magnetising current was then switched on, and the illumination in the field of the analyser was observed to increase or decrease according to the direction of the current. Within the limits of his experiments, which included incidences varying from 0° to 80°, Dr. Kerr found that the rotatory effect of reflection from magnetised iron was negative—that is, in a direction opposite to that of the magnetising current.

Kundt's experiments.

Subsequently Professor Kundt[1] extended the experiments to incidences ranging between 0° and 90°, and he found that the direction and amount of the rotatory effect depended not merely on the angle of incidence, but also on whether the plane of polarisation was parallel or perpendicular to the plane of incidence. When the

[1] Kundt, Berlin, *Sitzungsberichte*, July 10, 1884; *Phil. Mag.* vol. xviii. p. 308, 1884.

light was polarised in the plane of incidence the direction of rotation was found to be the same for all angles of incidence, and opposed to the Ampèrean molecular currents of the reflecting magnet. On the other hand, when the light was polarised at right angles to the plane of incidence, then the direction of rotation was the same as that of the Ampèrean current for angles of incidence between 0° and 80°, but in the opposite direction for incidences between 80° and 90°. In both cases the rotation reached a maximum at an incidence of about 65°. Further, it was found that iron exhibited an anomalous rotational dispersion, the red rays suffering a larger rotation than the blue by reflection from a magnet as well as by transmission through a magnetised film of iron.

Normal incidence. In order to examine the case of normal incidence Dr. Kerr found it necessary to modify the apparatus, as shown in Fig. 197. The wedge-shaped submagnet of Fig. 196 was dispensed with and replaced by a block of soft iron rounded at one end into the frustum of a cone. A small conical hole was drilled through the centre of this block, and it was then placed directly over the centre of the polar surface, with its axis vertical, as shown in Fig. 197, the block being separated from the pole of the magnet by a wide ring of writing-paper. Above the block a thin plate of glass C was placed, on which the horizontal beam of light from the polariser fell at an angle of 45°, and was reflected vertically downwards through the hole in the iron block. This beam, after suffering direct reflection at the surface of the pole, returned vertically, and passed in part through the plate C into the analyser A placed in position to receive it. The results obtained with this form of

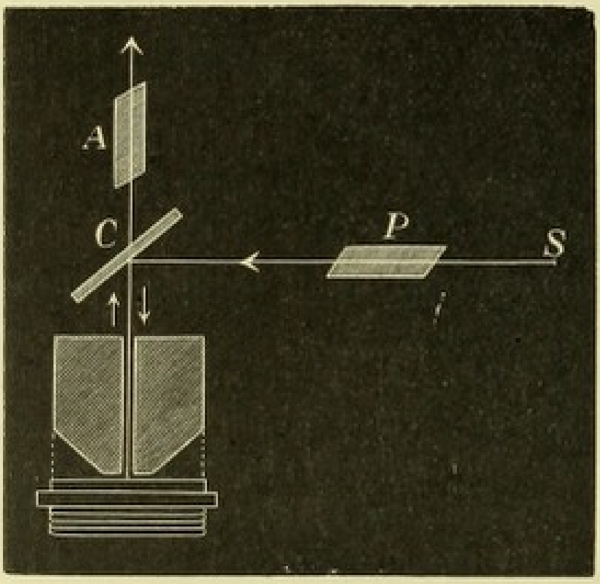


Fig. 197.

Glass plate. apparatus are complicated by the introduction of the glass plate C, for if any rotation occurs the light returning from the magnet to the plate will experience a further rotation during refraction through the plate, as explained in Art. 210. This disturbing effect of the plate seems to have been first recognised and allowed for by Professor Kundt.

In the foregoing experiments the lines of force are perpendicular to the reflecting surface, and to complete the investigation Dr. Kerr examined also the effect when the reflector is magnetised parallel to its surface. For this purpose the electromagnet was placed upright upon a table, and a rectangular bar of soft iron, one of whose sides was carefully planed and polished, was placed across the poles of the electromagnet, with the plane of its polished face vertical (the plane of the paper being horizontal), as shown in Fig. 198.

Force parallel to surface.

With this arrangement, in which the magnetic force is approximately parallel to the reflecting surface, it was found that when the light was polarised in the plane of incidence the magnetic rotation was negative—that is, opposite in direction to the Ampèrean current, which would produce the magnetisation—for all angles of incidence;

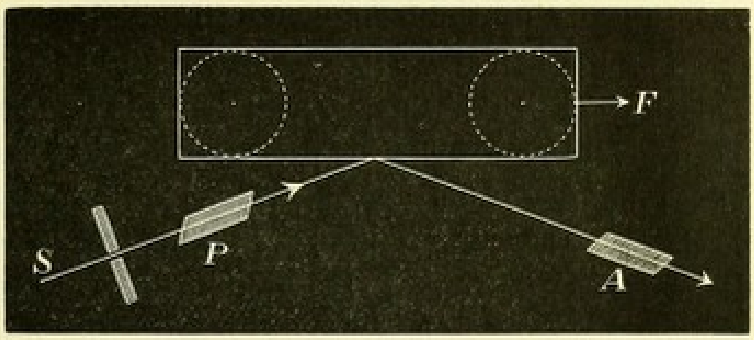


Fig. 198.

while for light polarised at right angles to the plane of incidence the rotation was negative for incidences between 90° and 75°, vanished at 75°, changed sign there, and remained positive for all angles of incidence between 75° and zero.

The particular case of perpendicular incidence on the side face was examined, and it was found that no optical effect was produced whether the plane of polarisation coincided with the direction of magnetisation or made an angle with it. From this it appears that there is no magneto-optic effect when the plane of wave front is parallel to the lines of magnetic force.

The optical effect produced when light is reflected at the surface of a magnetised substance may be brought into harmony with that produced by transmission through substances placed in a magnetic field, by the supposition that the light during reflection penetrates a thin film of the reflecting pole, and suffers rotation within the substance while traversing the film. This point of view has been worked out by Professor Kundt, who examined the light reflected from the second face of a glass plate. The faces of the plate were not parallel, but inclined at an angle, so that the light reflected from the first face was well separated from that which penetrated the plate, and, after suffering reflection at the second face, emerged again through the first. This twice refracted and once reflected beam was examined when the block of glass was placed upon the poles of an electromagnet with the lines of force parallel to the reflecting face, and it was found that when the incident light was polarised in the plane of incidence the emergent light was rotated in the positive direction for all angles of incidence, but that when the incident light was polarised at right angles to the plane of incidence the rotation was negative from normal incidence up to the polarising angle (56°·4), and was positive from the polarising angle to grazing incidence.

That a change of sign should occur at the polarising angle in the case of light polarised at right angles to the plane of incidence, and not in the case of light polarised in the plane of incidence, is indicated theoretically by the formulæ of Art. 210.

The mathematical investigation, and the explanation of the elliptic polarisation in these experiments on the basis of Maxwell's theory, have been given by Professor G. F. FitzGerald[1] and others. The rotatory effect discovered by Faraday when light passes along the lines of magnetic force, and the doubly refracting effect discovered by Kerr when light is transmitted across the lines of electric force, agree in this respect, that they both depend upon the presence of matter in the field; they both vary in sign according to the nature of the substance occupying the field, and they do not occur in the free ether.

**264. Division of a Plane-Polarised Ray into two Circularly Polarised Rays.**—The first theoretical interpretation of the phenomena of rotatory polarisation was given by Fresnel on the hypothesis that a rectilinear vibration may be regarded as the resultant of two opposite circular vibrations (see Art. 48, Ex. 2). According to this theory a plane-polarised ray may be regarded as the resultant of two opposite circularly polarised rays. If this be so a plane-polarised ray in passing through any substance may be decomposed into two others circularly polarised in opposite senses, which may or may not be propagated with the same velocity, according to the nature of the substance. If they travel with the same velocity then at emergence they will combine again into a plane-polarised ray, polarised in the same plane as before transmission, but if the opposite circular vibrations travel with different velocities then at emergence one will have suffered a retardation relatively to the other with the effect, as we shall presently show, that the plane of polarisation of the emergent beam will be inclined to the primitive plane of polarisation, and the rotation will be accounted for.

To test this point, and place it beyond doubt that this decomposition in reality occurs, and is not a result of the mere manipulation of formulæ, Fresnel devised the following experiment. Several prisms of quartz, alternately right-handed and left-handed, were arranged in the shape of a parallelopiped so as to form an achromatic combina-

[1] G. F. FitzGerald, *Phil. Trans.* 1880, pt. ii. p. 691.

tion. Thus in Fig. 199 the prisms ABC and BED are right-handed while BCD and DEF are left-handed, and in all the axis is parallel to the faces AE and CF. Now if a plane-polarised ray passing through quartz in the direction of the axis is really divided into two opposite circular vibrations which travel with different velocities, then on reaching the interface BC the quicker ray in ABC will be the slower in BCD and *vice versâ*, and consequently at the face BC one ray will be deviated from the normal and the other towards it. There will thus be a separation of the rays which will be augmented at the faces BD and DE, so that by sufficiently increasing the number of prisms the separation should be rendered visible. Fresnel found that this was the case, that two emergent pencils were presented, and that both were circularly polarised.[1]

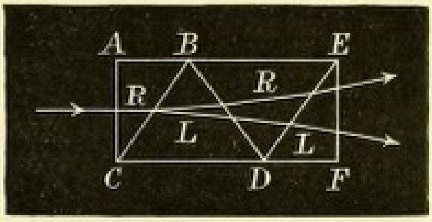



Fig. 199.

**265. Rotation of the Plane of Polarisation by Retardation of a Circular Component.**—Having shown that a plane-polarised ray may be decomposed into two opposite circularly polarised rays, it is easy to show that the effect of introducing a relative retardation between the circular vibrations is to rotate the plane of polarisation. Thus if two particles M and M′ (Fig 200) describe the same circle in opposite directions starting from the same point Y, it is clear that their velocities parallel to OY will be equal and in the same direction, while their velocities perpendicular to OY will be equal but oppositely directed. Hence if the two motions be simultaneously impressed on the same

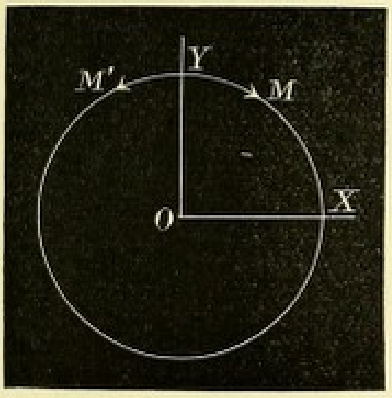



Fig. 200.

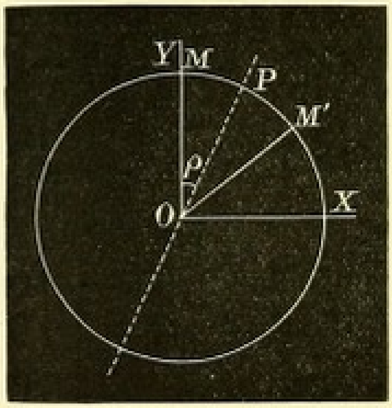



Fig. 201.

molecule its velocity parallel to OY, at any instant, will be double that of M or M′, and its velocity perpendicular to OY will be zero, so that

[1] The same results follow with ordinary light. In fact, the experiment was first devised to show the double refraction of unpolarised light (Fresnel, *Œuvres*, tom. i. p. 743, etc.).

it will simply vibrate backwards and forwards along the line OY with an amplitude of excursion equal to twice the radius of the circle.

Now let us suppose that one of the circular vibrations is, by any cause, relatively retarded so that when M travels round the circle to Y the other has not yet reached it, but has attained a position M′ (Fig. 201) at an angle $\delta$ from M. Draw OP bisecting the angle MOM′. If the two circular vibrations now move with the same velocity (which will be the case when the light emerges from the active substance) the two particles will cross at P, and their combined motions impressed on a single molecule will cause it to vibrate along the line OP. Thus if the original vibration in the incident plane-polarised light be parallel to OY, and if one of its circular components be retarded relatively to the other, by transmission through quartz, or otherwise, the plane of polarisation of the emergent light will be inclined to that of the incident light at an angle $\rho$ given by the equation

$$\rho = \tfrac{1}{2}\delta,$$

where $\delta$ is the relative phase retardation.

The same result may be arrived at very easily by means of the equations of the vibration. The incident vibration $y = 2a \sin \omega t$ is equivalent to the two opposite circular vibrations

$$x_1 = a \cos \omega t, \qquad y_1 = a \sin \omega t \qquad \text{(right-handed)},$$
$$x_2 = -a \cos \omega t, \qquad y_2 = a \sin \omega t \qquad \text{(left-handed)},$$

and if the latter be retarded in phase by an amount $\delta$ the equations of the transmitted vibrations may be written in the form

$$x_1 = a \cos \omega t, \qquad y_1 = a \sin \omega t,$$
$$x_2 = -a \cos (\omega t + \delta), \qquad y_2 = a \sin (\omega t + \delta).$$

Consequently we have for the transmitted vibration

$$x = a \cos \omega t - a \cos (\omega t + \delta) = 2a \sin \tfrac{1}{2}\delta \sin (\omega t + \tfrac{1}{2}\delta),$$
$$y = a \sin \omega t + a \sin (\omega t + \delta) = 2a \cos \tfrac{1}{2}\delta \sin (\omega t + \tfrac{1}{2}\delta).$$

These are two perpendicular vibrations of the same phase, which therefore compound into a resultant rectilinear vibration, making an angle $\rho$ with the axis OY, which is determined by the equation

$$\tan \rho = x/y = \tan \tfrac{1}{2}\delta.$$

Therefore, as before, $\rho = \tfrac{1}{2}\delta$.

An expression for the rotation $\rho$ may be easily found in terms of the wave lengths of the circular components. For if $\lambda_1$ and $\lambda_2$ be the wave lengths of the circular components in the quartz, then if $e$ be the thickness of the crystalline plate we have

$$e = n_1\lambda_1 = n_2\lambda_2,$$

where $n_1$ and $n_2$ are the numbers of vibrations executed by the two components during transmission. The relative phase retardation of the two components on emerging from the plate is consequently

$$\delta = 2\pi(n_1 - n_2) = 2\pi e\left(\frac{1}{\lambda_1} - \frac{1}{\lambda_2}\right).$$

Hence the expression for the rotation is

$$\rho = \pi e\left(\frac{1}{\lambda_1} - \frac{1}{\lambda_2}\right).$$

In the case of Iceland spar and other uniaxal crystals both rays travel with the same velocity in the direction of the axis. In such substances we have $\lambda_1 = \lambda_2$, and no rotation of the plane of polarisation is produced. In quartz, on the other hand, the wave surface consists, as is usually assumed, of a sphere and a concentric spheroid, but they do not touch. The radius of the sphere is greater than the major axis of the spheroid, so that the latter is entirely enclosed within the former and $\lambda_1$ is not equal to $\lambda_2$. Wave surface in quartz.

Experiment shows that $\rho$ is approximately proportional[1] to $1/\lambda^2$, we therefore conclude that $1/\lambda_1 - 1/\lambda_2$ is approximately proportional to $1/\lambda^2$. The formula for the rotation may then be written in the form

$$\rho = \frac{\kappa e}{\lambda^2},$$

where $e$ is the thickness of the substance passed through, $\lambda$ the wave length of the light in air, and $\kappa$ a constant depending on the nature of the substance.

**266. Decomposition of a Rectilinear Vibration into two opposite Elliptic Vibrations—Oblique Transmission through Quartz.**—It has been seen in the foregoing articles that two circular vibrations of the same amplitude and opposite senses compound into a rectilinear vibration, and conversely, it has been proved by experiment that a pencil of plane-polarised light is actually divided into two opposite circularly polarised beams by transmission through quartz in a direction parallel to the optic axis of the crystal.

From a mathematical point of view, however, the circular components have no particular claim to be regarded as the only legitimate constituents of a plane-polarised ray. A rectilinear vibration may be the resultant of a variety of systems of constituents, which may be rectilinear, or elliptic, or other than elliptic, instead of circular. Thus the rectilinear vibration

$$x = a \cos \omega t$$

[1] In fact, if $\lambda_2 = \lambda_1 + \delta\lambda_1$, $1/\lambda_1 - 1/\lambda_2 = \delta\lambda_1/\lambda_1^2$.

is mathematically equivalent to the two opposite elliptic vibrations

$$x_1=(a-k)\cos\omega t, \qquad y_1=b\sin(\omega t+\delta) \qquad \text{(right-handed)},$$
$$x_2=k\cos\omega t, \qquad y_2=-b\sin(\omega t+\delta) \qquad \text{(left-handed)},$$

for these, when added together, yield the original vibration. The choice of components must therefore be settled by experiment, and Fresnel's experiment (Art. 264) marks the circular components as the proper constituents when light is transmitted through quartz in the direction of the axis.

In directions inclined to the axis of the crystal, however, it was shown in 1831 by Sir G. B. Airy[1] that the decomposition must be regarded as elliptic, the ellipse becoming a circle in the particular case of transmission parallel to the optic axis. Thus when a plane-polarised ray is transmitted through rock-crystal in a direction inclined to the axis, it is divided into two others which are elliptically polarised; the elliptic vibrations in the two being in opposite directions, and their greater axes coinciding in direction with the principal plane and the perpendicular plane respectively. Thus the major axis of one elliptic vibration coincides in direction with the minor axis of the other, as in Fig. 202. The ratio of the axes is the same in both, but varies with the inclination of the ray to the optic axis, being equal to unity parallel to the optic axis, and to zero for directions perpendicular to it — the elliptic vibration including the circle and right line as particular cases. Thus a circularly polarised ray may pass through quartz in the direction of the optic axis without alteration, a plane-polarised ray in the perpendicular direction, and a given elliptic ray in some intermediate direction.

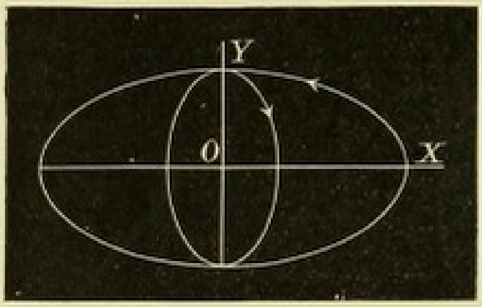


Fig. 202.

The phenomena presented by oblique transmission through quartz have all been explained satisfactorily by Airy on the assumption of this special decomposition, together with the supposition that one of the components is retarded relatively to the other during transmission.

In order to express this conveniently we shall take the initial rectilinear vibration to be of the form

$$x=(1+k^2)\cos\omega t.$$

It is evident that this may be replaced by the components

$$x_1=\cos\omega t, \qquad y_1=k\sin\omega t \qquad \text{(right-handed)},$$
$$x_2=k^2\cos\omega t, \qquad y_2=-k\sin\omega t \qquad \text{(left-handed)},$$

[1] G. B. Airy, *Trans. Cambridge Phil. Soc.* vol. iv. pp. 77, 199, 1833.

and these components are of the form required by Airy's theory, for they are elliptic, of opposite signs, and such that the major axis of one coincides with the minor axis of the other while the ratio of the axes is the same in both.

Now if the components $(x_2, y_2)$ be retarded by an amount $\delta$ relatively to the components $(x_1, y_1)$, the equations of the emergent vibration will be

$$x_1 = \cos\omega t, \qquad y_1 = k\sin\omega t,$$
$$x_2 = k^2\cos(\omega t + \delta), \qquad y_2 = -k\sin(\omega t + \delta).$$

Therefore for their resultant we have

$$x = \cos\omega t + k^2\cos(\omega t + \delta) = \mathrm{A}\cos(\omega t + \phi),$$

where

$$\mathrm{A}^2 = 1 + 2k^2\cos\delta + k^4, \quad \text{and } \tan\phi = \frac{k^2\sin\delta}{1 + k^2\cos\delta}.$$

Similarly

$$y = k\sin\omega t - k\sin(\omega t + \delta) = \mathrm{B}\cos(\omega t + \psi),$$

where

$$\mathrm{B}^2 = 4k^2\sin^2\tfrac{1}{2}\delta, \quad \text{and } \tan\psi = \tan\tfrac{1}{2}\delta.$$

If $k = 1$ the vibrations are circular, and we have $\phi = \psi = \frac{1}{2}\delta$. This applies to the case of propagation parallel to the axis.

**267. General Formula.**—When convergent plane-polarised light is transmitted through a thin plate of quartz cut perpendicularly to the axis, and then examined by means of an analyser, the isochromatic curves obtained exhibit some remarkable peculiarities which are not presented in the case of crystals which do not possess rotatory power. To determine these peculiarities it becomes necessary to obtain an expression for the intensity in the field of the analyser when the light traverses the quartz plate in a direction oblique to the axis.

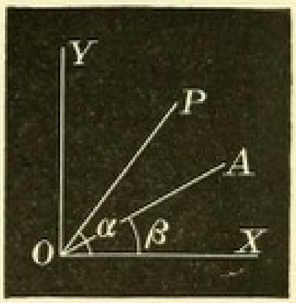


Fig. 203.

Let the incident light be plane-polarised, and let OP (Fig. 203) be parallel to the principal plane of the polariser, OX and OY to the directions of vibration in the crystal. Then if the angle POX $= \alpha$, the incident vibration, being parallel to OP by Fresnel's theory, will give components of amplitudes $a\cos\alpha$ and $a\sin\alpha$ parallel to OX and OY respectively.

Writing $a = 1 + k^2$, the vibration $x = (1 + k^2)\cos\alpha\cos\omega t$ parallel to OX may be replaced by the two opposite elliptic vibrations

$$\left.\begin{array}{lll}\text{(right-handed)} & x_1 = \cos\alpha\cos\omega t, & y_1 = k\cos\alpha\sin\omega t \\ \text{(left-handed)} & x_2 = k^2\cos\alpha\cos\omega t, & y_2 = -k\cos\alpha\sin\omega t\end{array}\right\} \quad (1),$$

in which the ratio of the axes of one is equal to the ratio of the axes of the other, but the minor axis of one is in the direction of the major axis of the other, as indicated in Fig. 202. So also the vibration $y=(1+k^2)\sin\alpha\cos\omega t$ parallel to OY may be replaced by the elliptic vibrations

$$\begin{array}{lll} \text{(left-handed)} & x'_1 = k\sin\alpha\sin\omega t, & y'_1 = \sin\alpha\cos\omega t \\ \text{(right-handed)} & x'_2 = -k\sin\alpha\sin\omega t, & y'_2 = k^2\sin\alpha\cos\omega t \end{array} \Bigg\} \quad (2).$$

In passing through the crystal one (the left-handed, say) of the elliptic vibrations is retarded on the other by an amount $\delta$, so that on emergence the vibrations (1) are of the form

$$\begin{array}{ll} x_1 = \cos\alpha\cos\omega t, & y_1 = k\cos\alpha\sin\omega t \\ x_2 = k^2\cos\alpha\cos(\omega t-\delta), & y_2 = -k\cos\alpha\sin(\omega t-\delta) \end{array} \Bigg\} \quad (3),$$

and equations (2) become

$$\begin{array}{ll} x'_1 = k\sin\alpha\sin(\omega t-\delta), & y'_1 = \sin\alpha\cos(\omega t-\delta) \\ x'_2 = -k\sin\alpha\sin\omega t, & y'_2 = k^2\sin\alpha\cos\omega t \end{array} \Bigg\} \quad (4).$$

Hence if the principal plane of the analyser OA makes an angle $\beta$ with OX, we have for the vibration in that plane

$$Y = (x_1+x_2+x'_1+x'_2)\cos\beta + (y_1+y_2+y'_1+y'_2)\sin\beta,$$

from which the general expression for the intensity in the field of the analyser may be obtained. Let us first consider the case in which the Nicols are crossed.

**268. Nicols Crossed.**—In the particular case in which the polariser and analyser are crossed we have $\beta-\alpha=90°$, and the calculation becomes much simplified. Thus the vibration in the principal plane of the analyser is

$$\begin{aligned} Y &= (y_1+y_2+y'_1+y'_2)\cos\alpha - (x_1+x_2+x'_1+x'_2)\sin\alpha, \\ &= 2k\sin\tfrac{1}{2}\delta\cos(\omega t-\tfrac{1}{2}\delta) + (1-k^2)\sin 2\alpha\sin\tfrac{1}{2}\delta\sin(\omega t-\tfrac{1}{2}\delta). \end{aligned}$$

Hence the intensity is given by the equation

$$I = \{4k^2 + (1-k^2)^2\sin^2 2\alpha\}\sin^2\tfrac{1}{2}\delta \quad (1).$$

When the incident light is a parallel beam, and the thickness of the plate uniform, the directions of vibration OX and OY will be the same at all points of the plate, so that $\alpha$ remains constant as well as $\delta$. The colour and intensity will therefore be the same at all points unless the plate possesses irregular structure.

With a convergent beam on the other hand $\delta$ will increase as we proceed along any radius vector drawn from the centre of the field. Further, the two directions of vibration within the plate for a ray incident at any point X will be along and perpendicular to OX, as explained in Art. 232, so that these directions vary as X rotates around the centre O. The angle $\alpha$ will consequently not be constant, and in the foregoing formula for the intensity it will play the part of the direction angle to the radius vector drawn to the point X, at which the intensity is expressed. Hence, in general, as X moves over the face of the crystal the intensity will vary by reason of the variations of $\alpha$ as well as those of $\delta$. For directions slightly inclined to the axis, $k=1$, and we have approximately

$$I = 4 \sin^2 \tfrac{1}{2}\delta \qquad (2).$$

Consequently in the neighbourhood of the optic axis the intensity depends only on the value of $\delta$, and as $\delta$ will be different for the different colours, it follows that when white light is used the central spot will be coloured and will not be crossed by any uncoloured lines. Central spot.

In directions considerably inclined to the axis, we have approximately $k=0$, and the expression for the intensity becomes

$$I = \sin^2 2\alpha \sin^2 \tfrac{1}{2}\delta \qquad (3).$$

The term $\sin 2\alpha$ gives rise to a dark rectangular cross (Fig. 204) corresponding to $\alpha = 0°$ and $\alpha = 90°$, while the term $\sin^2 \frac{1}{2}\delta$ produces a system of concentric circular rings, as in the case of an ordinary uniaxal crystal, alternately dark and bright when monochromatic light is used, but with white light they will be iris-coloured. The pattern presented in the field consequently consists of a cross, together with a system of coloured rings surrounding a central spot of tint depending on the thickness of the plate. Cross and rings.

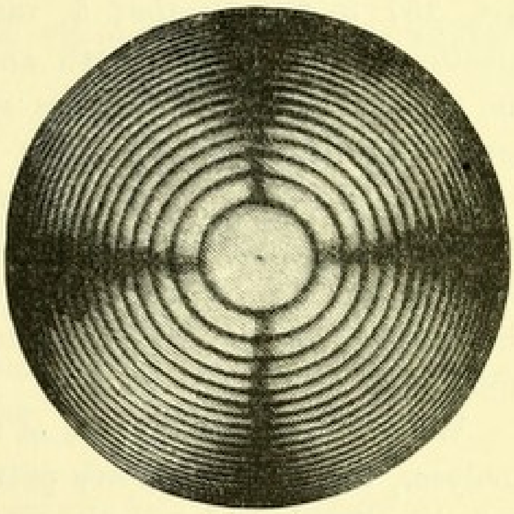


Fig. 204.

It should be observed that the intensity, as expressed by formula (1), will vanish only when $\delta$ is zero, or a multiple of $\frac{1}{2}\pi$, and for this reason the dark cross of the pattern is not black. The illumination is simply least along the directions $\alpha = 0$ and $\alpha = \frac{1}{2}\pi$, etc., and greatest along the lines $\alpha = \frac{1}{4}\pi$ and $\alpha = \frac{3}{4}\pi$, etc. Near the centre of the field $k$ is sensibly equal to unity, and the intensity given by equation (2) is the same in all directions around the centre. The dark cross consequently does not begin to appear until we recede to some distance from the centre, and its arms do not completely interrupt the rings in any region. They merely approach blackness at a considerable distance from the centre where $k$ is sensibly equal to zero, and where the intensity is given by equation (3).

In the case of uniaxal non-rotating crystals the cross is black and traverses the centre of the field, as shown in Fig. 182. This happens because the ordinary and extraordinary waves travel with the same velocity in the direction of the optic axis, and consequently $\delta$ is zero at the centre of the field. In the case of rotating crystals, however, the two waves travel with different velocities, even in the direction of the optic axis. According to Airy's theory the two sheets of the wave surface do not touch each other in a rotating crystal. The radius of the sphere in quartz is greater than the greatest axis of the spheroid, so that the latter is contained within the former, and as a consequence $\delta$ is not zero at the centre of the field. Now in the case of non-rotating crystal the expression for $\delta$ at a distance $r$ from the centre is (p. 395) of the form $\delta = mr^2$, where $m$ is a constant depending on the nature of the plate. Hence in the case of a rotating plate the retardation is of the form

Rotating and non-rotating crystals compared.

$$\delta = mr^2 + \delta_0,$$

$\delta_0$ being the retardation at the centre of the field where $r = 0$. By Art. 257 it follows that $\delta_0$ varies inversely as $\lambda^2$, so that $\delta$ is a function of the wave length and will not vanish simultaneously for the various colours. For this reason when white light is used the central spot is never black but coloured, with a tint which depends upon the thickness of the plate. As in the case of a non-rotating plate the difference between the squares of the radii of two consecutive rings is constant for light of a given wave length, but the squares of the radii are no longer proportional to the natural numbers.

When the principal plane of the polariser is parallel to that of the analyser, the complementary pattern is exhibited. In this case $\beta = \alpha$, and the vibration is

$$Y = (x_1 + x_2 + x'_1 + x'_2)\cos\alpha + (y_1 + y_2 + y'_1 + y'_2)\sin\alpha.$$

Hence we find

$$I' = (1 + k^2)^2 - \{4k^2 + (1 - k^2)^2 \sin^2 2\alpha\} \sin^2 \tfrac{1}{2}\delta,$$

an expression which might have been written down from the value of I with the consideration that I and I′ must be complementary, and that the amplitude of the incident light was taken equal to $(1 + k^2)$.

If a doubly refracting rhomb be used as analyser, I and I′ represent the intensities of the extraordinary and ordinary images respectively.

**269. General Expression for the Intensity.**[1]—When the principal planes of the Nicols are inclined to each other, the calculation of the intensity becomes more complicated.

Substituting from equations (3) and (4) of Art. 267 we have

$$\begin{aligned} \mathrm{Y} = \cos\beta\,\{&\cos\omega t\,(\cos\alpha + k^2\cos\alpha\cos\delta - k\sin\alpha\sin\delta) \\ &+ \sin\omega t\,(k^2\cos\alpha\sin\delta - k\sin\alpha + k\sin\alpha\cos\delta)\} \\ + \sin\beta\,\{&\cos\omega t\,(k\cos\alpha\sin\delta + k^2\sin\alpha + \sin\alpha\cos\delta) \\ &+ \sin\omega t\,(k\cos\alpha - k\cos\alpha\cos\delta + \sin\alpha\sin\delta)\}. \end{aligned}$$

Writing this in the form

$$\mathrm{Y} = \mathrm{A}\cos\omega t + \mathrm{B}\sin\omega t,$$

the intensity is given by the equation

$$\mathrm{I} = \mathrm{A}^2 + \mathrm{B}^2,$$

where

$$\begin{aligned} \mathrm{A} &= \cos\beta\,(\cos\alpha + k^2\cos\alpha\cos\delta - k\sin\alpha\sin\delta) \\ &\qquad + \sin\beta\,(k\cos\alpha\sin\delta + k^2\sin\alpha + \sin\alpha\cos\delta), \\ &= k^2(\sin\alpha\sin\beta + \cos\alpha\cos\beta\cos\delta) - k\sin(\alpha-\beta)\sin\delta \\ &\qquad + \cos\alpha\cos\beta + \sin\alpha\sin\beta\cos\delta, \end{aligned}$$

and

$$\begin{aligned} \mathrm{B} &= \cos\beta\,(k^2\cos\alpha\sin\delta - 2k\sin\alpha\sin^2\tfrac{1}{2}\delta) \\ &\qquad + \sin\beta\,(\sin\alpha\sin\delta + 2k\cos\alpha\sin^2\tfrac{1}{2}\delta), \\ &= k^2\cos\alpha\cos\beta\sin\delta - 2k\sin(\alpha-\beta)\sin^2\tfrac{1}{2}\delta + \sin\alpha\sin\beta\sin\delta. \end{aligned}$$

Squaring and adding these expressions, we find that I may be written in the form

$$\mathrm{I} = \mathrm{L}k^4 - \mathrm{M}k^3 + \mathrm{N}k^2 - \mathrm{M}k + \mathrm{L}.$$

The coefficient of $k^4$ is

$$\begin{aligned} \mathrm{L} &= (\sin\alpha\sin\beta + \cos\alpha\cos\beta\cos\delta)^2 + \cos^2\alpha\cos^2\beta\sin^2\delta, \\ &= \sin^2\alpha\sin^2\beta + \cos^2\alpha\cos^2\beta + 2\sin\alpha\cos\alpha\sin\beta\cos\beta\cos\delta, \\ &= \cos^2(\alpha-\beta) - \sin 2\alpha\sin 2\beta\sin^2\tfrac{1}{2}\delta, \end{aligned}$$

and the term independent of $k$ is at once seen to have the same value.

Seeking the coefficient of $k^3$ we have

$$\begin{aligned} \mathrm{M} &= 2(\sin\alpha\sin\beta + \cos\alpha\cos\beta\cos\delta)\sin(\alpha-\beta)\sin\delta \\ &\qquad + 4\cos\alpha\cos\beta\sin(\alpha-\beta)\sin\delta\sin^2\tfrac{1}{2}\delta, \\ &= 2(\sin\alpha\sin\beta + \cos\alpha\cos\beta\cos\delta + 2\cos\alpha\cos\beta\sin^2\tfrac{1}{2}\delta)\sin(\alpha-\beta)\sin\delta, \\ &= 2\sin(\alpha-\beta)\cos(\alpha-\beta)\sin\delta, \\ &= \sin 2(\alpha-\beta)\sin\delta, \end{aligned}$$

and we obtain the same expression for the coefficient of $k$.

Finally the coefficient of $k^2$ is

$$\begin{aligned} \mathrm{N} &= 2(\sin\alpha\sin\beta + \cos\alpha\cos\beta\cos\delta)(\cos\alpha\cos\beta + \sin\alpha\sin\beta\cos\delta) \\ &\qquad + \sin^2(\alpha-\beta)\sin^2\delta + 2\sin\alpha\cos\alpha\sin\beta\cos\beta\sin^2\delta + 4\sin^2(\alpha-\beta)\sin^4\tfrac{1}{2}\delta, \\ &= 2\cos\delta\,(\sin^2\alpha\sin^2\beta + \cos^2\alpha\cos^2\beta) + \sin 2\alpha\sin 2\beta + 4\sin^2(\alpha-\beta)\sin^2\tfrac{1}{2}\delta, \\ &= 2\cos^2(\alpha-\beta) - 4\sin^2\tfrac{1}{2}\delta\,\{\sin^2\alpha\sin^2\beta + \cos^2\alpha\cos^2\beta - \sin^2(\alpha-\beta)\}, \\ &= 2\cos^2(\alpha-\beta) - \sin^2\tfrac{1}{2}\delta\,\{4\cos 2(\alpha-\beta) - 2\sin 2\alpha\sin 2\beta\}. \end{aligned}$$

[1] See note by G. Quesneville, *Comptes Rendus*, cxxi. pp. 522-524, 1885.

Hence

$$\begin{aligned} \mathrm{I} &= (k^4+1)\{\cos^2(\alpha-\beta) - \sin 2\alpha \sin 2\beta \sin^2 \tfrac{1}{2}\delta\} - k(1+k^2)\sin 2(\alpha-\beta)\sin\delta \\ &\qquad + k^2[2\cos^2(\alpha-\beta) - \{4\cos 2(\alpha-\beta) - 2\sin 2\alpha \sin 2\beta\}\sin^2\tfrac{1}{2}\delta], \\ &= (1+k^2)^2\cos^2(\alpha-\beta) - k(1+k^2)\sin 2(\alpha-\beta)\sin\delta \\ &\qquad - \{4k^2\cos 2(\alpha-\beta) + (1-k^2)^2 \sin 2\alpha \sin 2\beta\}\sin^2\tfrac{1}{2}\delta. \end{aligned}$$

Which may also be written in the form

$$\mathrm{I} = [(1+k^2)\cos(\alpha-\beta)\cos\tfrac{1}{2}\delta - 2k\sin(\alpha-\beta)\sin\tfrac{1}{2}\delta]^2 + (1-k^2)^2\cos^2(\alpha+\beta)\sin^2\tfrac{1}{2}\delta.$$

*Cor.* 1.—If $k=0$, we have

$$\mathrm{I} = \cos^2(\alpha-\beta) - \sin 2\alpha \sin 2\beta \sin^2\tfrac{1}{2}\delta,$$

which is the ordinary formula for uniaxal crystals devoid of rotatory power.

*Cor.* 2.—If $k=1$—that is, if the ray passes along the axis—we have

$$\mathrm{I} = 4\{\cos(\alpha-\beta)\cos\tfrac{1}{2}\delta - \sin(\alpha-\beta)\sin\tfrac{1}{2}\delta\}^2 = 4\cos^2(\alpha-\beta+\tfrac{1}{2}\delta).$$

This expression is a maximum when $\beta-\alpha=\frac{1}{2}\delta$, and zero when $\alpha-\beta = \frac{1}{2}(\pi-\delta)$. This indicates that the light emerging from the quartz plate is plane-polarised, and that the plane of polarisation has been rotated through an angle $\beta-\alpha=\frac{1}{2}\delta$.

*Cor.* 3.—When the Nicols are crossed $\alpha-\beta=90^\circ$, and, as in Art. 268, we have

$$\mathrm{I} = \{4k^2 + (1-k^2)^2\sin^2 2\alpha\}\sin^2\tfrac{1}{2}\delta.$$

**270. Approximate Form of the Isochromatic Lines.**—When the incident light is parallel the angles $\alpha$ and $\beta$ have the same values at every point of the plate, and the intensity is uniform. With a convergent beam, however, the angle $\alpha$ is variable, and, as already remarked, is the direction angle of the radius vector to the point under consideration. For given positions of the Nicols the angle $\alpha-\beta$ will be constant, viz. the angle between their principal planes, and denoting it by $\gamma$, the expression for the intensity may be written in the form

$$\mathrm{I} = \{(1+k^2)\cos\gamma\cos\tfrac{1}{2}\delta - 2k\sin\gamma\sin\tfrac{1}{2}\delta\}^2 + (1-k^2)\cos^2(2\alpha-\gamma)\sin^2\tfrac{1}{2}\delta.$$

In this expression $\delta$ and $k$ vary from point to point, and at any given point $\delta$ is a function of the wave length, so that the investigation of the isochromatic lines becomes very complicated. An idea of their approximate form may, however, be obtained by finding the shape of the bright and dark lines for the case of monochromatic light, on the supposition that along any one of these lines $k$ is sensibly constant—

that is, that the different points of a bright or dark line are not at very different distances from the centre. In accordance with this supposition we have merely to express that I is a maximum or a minimum when $\delta$ is regarded as the variable. For this purpose Airy made use of a subsidiary angle $\chi$, determined by the equation

$$\tan\chi = \frac{2k}{1+k^2}\tan\gamma.$$

When this is substituted in the first term of the expression for I it becomes

$$(1+k^2)^2\cos^2\gamma\sec^2\chi\cos^2(\chi+\tfrac{1}{2}\delta),$$

or, replacing $\sec^2\chi$, the expression for I takes the form

$$\mathrm{I} = \{(1+k^2)^2\cos^2\gamma + 4k^2\sin^2\gamma\}\cos^2(\chi+\tfrac{1}{2}\delta) + (1-k^2)^2\cos^2(2\alpha-\gamma)\sin^2\tfrac{1}{2}\delta.$$

The points of maximum and minimum intensity along the radius vector corresponding to a given value of $\alpha$ are found from the equation $d\mathrm{I}/d\delta = 0$. Writing the expression for I in the form

$$\mathrm{I} = \mathrm{P}\cos^2(\chi+\tfrac{1}{2}\delta) + \mathrm{Q}\sin^2\tfrac{1}{2}\delta,$$

and differentiating with regard to $\delta$, we have

$$\mathrm{P}\sin(2\chi+\delta) = \mathrm{Q}\sin\delta.$$

From this equation it follows immediately that

$$\frac{\sin(2\chi+\delta)+\sin\delta}{\sin(2\chi+\delta)-\sin\delta} = -\frac{\mathrm{P}+\mathrm{Q}}{\mathrm{P}-\mathrm{Q}}.$$

Hence

$$\tan(\chi+\delta) = -\tan\chi\frac{\mathrm{P}+\mathrm{Q}}{\mathrm{P}-\mathrm{Q}},$$

or substituting for P and Q we have finally

$$\tan(\chi+\delta) = -\tan\chi\frac{(1+k^2)^2\cos^2\gamma + 4k^2\sin^2\gamma + (1-k^2)^2\cos^2(2\alpha-\gamma)}{(1+k^2)^2\cos^2\gamma + 4k^2\sin^2\gamma - (1-k^2)^2\cos^2(2\alpha-\gamma)}.$$

Writing $-\tan\Omega$ for the right-hand member of this equation we have

$$\tan(\chi+\delta) = -\tan\Omega,$$

and consequently

$$\chi+\delta = \pi-\Omega,$$

where $\Omega$ is an angle which varies with $\alpha$ as the point under consideration moves along a bright or dark line. The polar equation of any such line may be expressed by substituting $cr^2+\delta_0$ for $\delta$ in the foregoing (as in Art. 268), and we then have

$$cr^2 = (\pi-\Omega-\chi)-\delta_0,$$

where $\chi$ is a constant, when $\gamma$ and $k$ are given, and $\Omega$ varies with $a$ only. Hence, as the point moves along a bright or dark band, the radius vector $r$ will increase or decrease according as $\Omega$ decreases or increases. The rings will consequently deviate from the circular form by amounts depending on the variations of $\Omega$. Now $\Omega$ is greatest when $2a - \gamma = n\pi$—that is, in the four rectangular positions of the radius vector corresponding to

$$a = \tfrac{1}{2}\gamma, \quad a = \tfrac{1}{2}(\gamma + \pi), \quad a = \tfrac{1}{2}(\gamma + 2\pi), \quad a = \tfrac{1}{2}(\gamma + 3\pi).$$

But $\gamma = a - \beta$, and consequently the equation $a = \frac{1}{2}\gamma$ is the same as $a = -\beta$, and this means that the radius vector of a bright or dark band will be least along a pair of rectangular lines such that at each point of them the angle $\gamma$ between the principal planes of the polariser

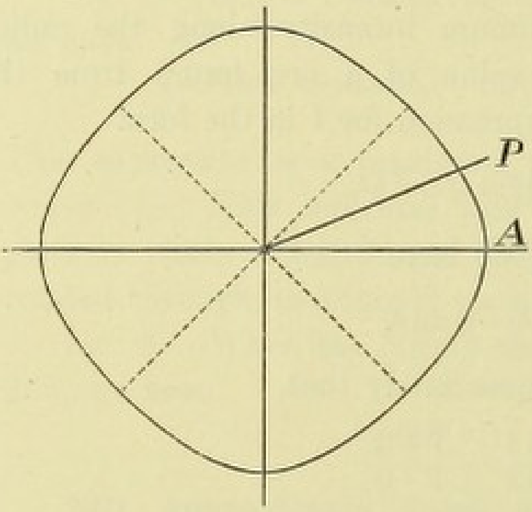



Fig. 205.

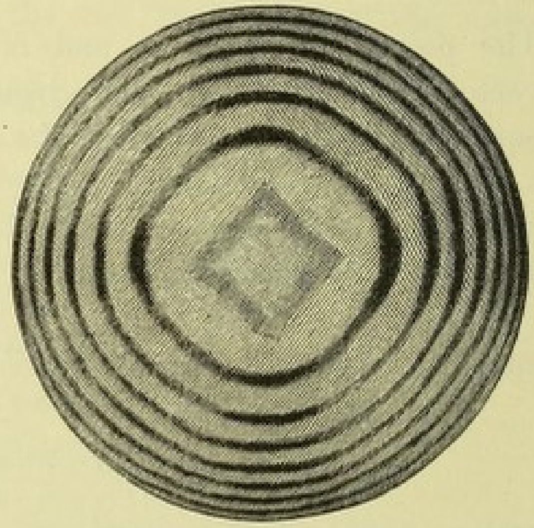


Fig. 206.

and analyser is bisected by the principal planes of the plate. The greatest values of $r$ correspond to the least values of $\Omega$, and these occur when $2a - \gamma = (n + \frac{1}{2})\pi$—that is, in the four rectangular positions

$$a = \tfrac{1}{2}(\gamma + \tfrac{1}{2}\pi), \quad a = \tfrac{1}{2}(\gamma + \tfrac{3}{2}\pi), \text{ etc.}$$

Hence the positions of the maxima differ from those of the minima by $45°$, and the general shape of a bright or dark band is a square-shaped curve such as that shown in Fig. 205, the greatest values of the radius vector corresponding to the diagonals of the square. Close to the centre $k$ is nearly equal to unity, and the expression for I will vanish when $\chi + \frac{1}{2}\delta = (n + \frac{1}{2})\pi$. But at the centre $\chi = \gamma$, consequently when $\delta = (2n + 1)\pi - 2\gamma$, there will be a black spot at the centre. Now, for points at a given small distance from the centre $\delta$ is constant, and I will be greatest when $2a - \gamma = n\pi$, and least in the directions $2a - \gamma = (n + \frac{1}{2})\pi$—that is, along the diagonals of the square curve.

The spot at the centre will consequently be a cross whose arms are along the diagonals of the square. With white light this central cross will be coloured. At a considerable distance from the centre $k$ approaches zero, and $\Omega$ becomes very small, and remains sensibly constant along any band, so that in the outer regions the bands become sensibly circular, as shown in Fig. 206.

**271. Convergent Circularly Polarised Light — Calculation of the Intensity.**—When the incident light is circularly polarised we may represent the component vibrations parallel to the principal sections of the thin plate by

$$x=(1+k^2)\cos\omega t, \quad y=-(1+k^2)\sin\omega t,$$

if the amplitude of the incident light be $(1+k^2)\sqrt{2}$. These components become respectively, after transmission through the plate,

$$\begin{array}{lll} \text{(right-handed)} & x_1=\cos\omega t, & y_1=k\sin\omega t \\ \text{(left-handed)} & x_2=k^2\cos(\omega t-\delta), & y_2=-k\sin(\omega t-\delta) \end{array}\Bigg\},$$
$$\begin{array}{lll} \text{(left-handed)} & x'_1=k\cos(\omega t-\delta), & y'_1=-\sin(\omega t-\delta) \\ \text{(right-handed)} & x'_2=-k\cos\omega t, & y'_2=-k^2\sin\omega t \end{array}\Bigg\}.$$

Hence for the vibration parallel to the principal plane of the analyser we have

$$\begin{aligned} Y=\cos\beta\{\cos\omega t(1+k^2\cos\delta+k\cos\delta-k)+\sin\omega t(k^2\sin\delta+k\sin\delta)\} \\ +\sin\beta\{\cos\omega t(k\sin\delta+\sin\delta)+\sin\omega t(k-k\cos\delta-\cos\delta-k^2)\}. \end{aligned}$$

Denoting the coefficient of $\cos\omega t$ by A and that of $\sin\omega t$ by B we have $I=A^2+B^2$, where

$$\begin{aligned} A&=k^2\cos\beta\cos\delta+k(\cos\beta\cos\delta+\sin\beta\sin\delta-\cos\beta)+\cos\beta+\sin\beta\sin\delta, \\ B&=k^2(\cos\beta\sin\delta-\sin\beta)+k(\cos\beta\sin\delta-\sin\beta\cos\delta+\sin\beta)-\sin\beta\cos\delta, \end{aligned}$$

and the expression for I becomes

$$I=(1+k^2)^2-4k(1-k^2)\cos 2\beta\sin^2\tfrac{1}{2}\delta+(1-k^4)\sin 2\beta\sin\delta.$$

*Cor.* 1.—If $k=0$, we have, as in Art. 237,

$$I=1+\sin 2\beta\sin\delta.$$

*Cor.* 2.—If $k=1$, we have

$$I=4,$$

that is, the light is transmitted parallel to the axis without alteration.

**272. Isochromatic Lines—Airy's Spirals.**—In order to determine the form of the bright and dark lines in the case of circularly polarised light the general expression for I obtained in the preceding article may be transformed, as in Art. 270, by using a subsidiary angle $\chi$ related to $\beta$ by the equation

$$\tan\chi=\frac{1+k^2}{2k}\tan 2\beta.$$

Substituting this in the expression for I it becomes

$$I = (1 + k^2)^2 + 2(1 - k^4) \sin 2\beta \operatorname{cosec} \chi \sin \tfrac{1}{2}\delta \sin(\chi - \tfrac{1}{2}\delta)$$
$$= (1 + k^2)^2 + (1 - k^4) \sin 2\beta \operatorname{cosec} \chi \{\cos(\chi - \delta) - \cos \chi\}.$$

Now if we consider a radius vector corresponding to a given value of $\beta$, and if we regard $k$ as sensibly constant in passing along this radius from a bright band to the adjacent dark band, then $\chi$ will also be sensibly constant within this range, and I will pass from a maximum to a minimum when $\chi - \delta$ passes from $2n\pi$ to $(2n+1)\pi$. Hence, regarding $\chi$ as a function of $\beta$ as we pass along a dark band, we will have

$$\chi - \delta = (2n+1)\pi,$$

or writing $cr^2 + \delta_0$ for $\delta$ as before, we have for the equation of the band

$$cr^2 = \chi - (2n+1)\pi - \delta_0.$$

For a given value of $n$ this represents a spiral curve in which $r$ continually increases as $\chi$ increases positively. When $n$ is increased

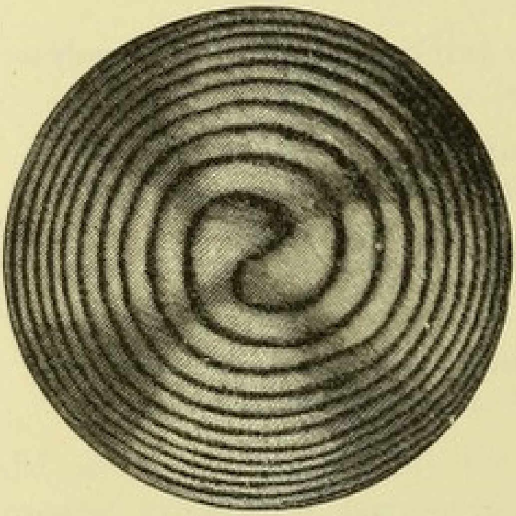


Fig. 207.

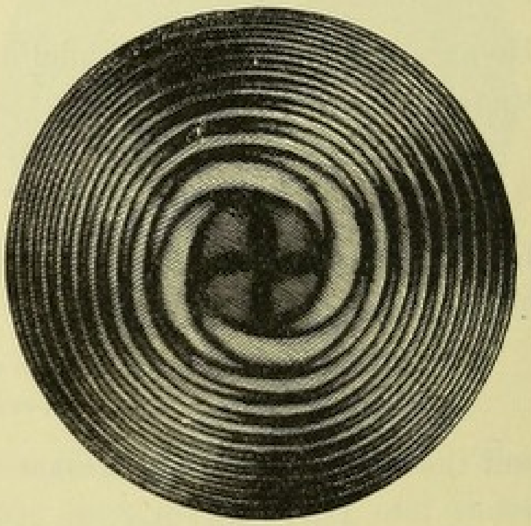


Fig. 208.

by unity a second spiral is obtained which is the same as the first, rotated through two right angles, and when $n$ is changed into $n+2$ the first spiral is repeated, and so on. The complete figure consequently consists of two similar spiral curves, mutually enwrapping each other, their positions differing by 180°, as shown in Fig. 207.

Close to the centre $k$ is sensibly equal to unity, and the intensity is approximately equal to $(1+k^2)^2$, so that the central tint is sensibly white. The spirals consequently will not appear very near the centre, and at a distance from the centre they will be faint in the directions parallel and perpendicular to the principal plane of the analyser—in other words, the outer regions will be traversed in these directions by a dark cross.

When the quartz plate is right-handed and the incident light right-handed the spirals enwrap each other from right to left, and in the reverse manner when the sign of the plate or that of the incident light is changed.[1]

## *Examples*

1. Plane-polarised light passes in succession through two quartz plates of equal thickness and opposite rotations; show that the intensity of the illumination in the analyser, when the Nicols are crossed, is

$$I=4(1-k^2)^2\left\{\frac{2k}{1+k^2}\cos 2\alpha\sin\tfrac{1}{2}\delta+\sin 2\alpha\cos\tfrac{1}{2}\delta\right\}^2\sin^2\tfrac{1}{2}\delta.$$

2. The curves of zero intensity in Ex. 1 are given by the equations

$$\sin\tfrac{1}{2}\delta=0,$$
$$2k\tan\tfrac{1}{2}\delta+(1+k^2)\tan 2\alpha=0.$$

The former gives a system of concentric circular rings, and the latter represents four distinct spirals passing through the points where the circles meet the axes OX and OY (Fig. 208).

## Applications

**273. Analysers.**—The peculiar property which some substances possess of rotating the plane of polarisation of light may be used as a means of detecting their presence in solutions when mixed with other substances which are inactive; and not only does it afford a qualitative test, but also an exact determination of the amount of active substance present, when we know the amount by which the plane of polarisation is rotated. The application of this property to the analysis of solutions consequently involves the exact measurement of the rotation of the plane of polarisation, and this involves primarily the accurate determination of the position of the plane of polarisation. This determination cannot be made with sufficient precision by means of an analyser such as a Nicol's prism alone, for with such an analyser the position of the plane of polarisation of a beam is determined by placing the Nicol so that the light is completely quenched, or else transmitted most copiously. Either of these positions can be attained with only moderate certainty, as it is difficult to ascertain when the field is exactly brightest or darkest, for near the position of maximum or minimum illumination the change of intensity, as the Nicol is rotated, is small, and is scarcely appreciable for a small displacement in either direction.

To remedy this defect analysers of special construction have been

[1] Coloured drawings of these figures are given in Airy's *Memoir*.

invented which involve the use of an additional piece of apparatus in conjunction with the analysing Nicol. This piece is so designed that the position of the plane of polarisation is determined by comparing the equality of two lights seen simultaneously and in juxtaposition. The superiority of this method depends upon the fact that we can say with much more certainty when two parts of the field are equally illuminated than when any part of it has attained its least or greatest brightness.

**274. Jellett's Analyser and Saccharimeter.**—The analyser devised by Jellett enables us to determine the position of the plane of polarisation by comparing the equality of two lights seen at the same time and in juxtaposition. To construct the additional piece a rhomb

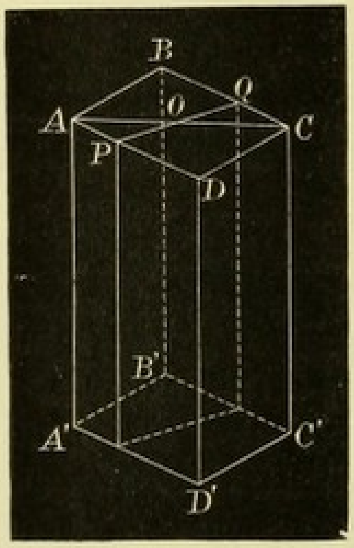



Fig. 209.

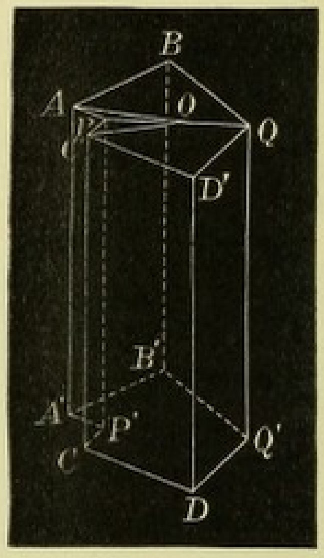



Fig. 210.

of Iceland spar, about two inches in length, is taken, and its ends are sawn off at right angles to its length, so as to form a right prism ABCD (Fig. 209). This prism is divided by a plane through PQ parallel to its edges, and making a small angle with the longer diagonal AC of the base ABCD. One of the two halves, PQCD, is now reversed, and the two are cemented together with their surfaces of section in contact as represented in Fig. 210, which is obtained from Fig. 209 by rotating PQCD through two right angles, keeping the two pieces in contact at their common surface of section.

Let us now suppose that a pencil of plane-polarised light falls normally on the face ABCD. The ordinary ray passes through undeviated and polarised in a plane perpendicular to AC—that is, according to Fresnel's hypothesis, vibrating parallel to AC—while the extraordinary ray is deviated to one side, and, if the rhomb be sufficiently long, the two will be separated at emergence. Hence in the Jellett prism ABQD′C′ (Fig. 210), when light is transmitted through it normally, the vibration of the ordinary ray will be parallel to AO in the part AOQB and parallel to C′O in the part C′OQD′. Consequently if a cylindrical pencil of light traverse the prism so as to be equally divided by the plane of section OQ, the extraordinary rays will be thrown off in opposite directions in the two halves, and may be stopped by a diaphragm suitably placed. The ordinary rays, on the other hand, will be transmitted without deviation, and will constitute a cylindrical pencil equally divided by the plane of section, but polarised in planes inclined at an angle $2\alpha = \text{AOC}'$, where $\alpha = \text{AOP}$, the angle which the plane of section makes with the diagonal AC (Fig. 209).

Now if the incident light be plane-polarised, the component parallel to AO is transmitted by one half, and the component parallel to C′O by the other, and therefore if we look through the analyser, the intensities in the two halves of the field will be different, except when the plane of the section is parallel to the primitive plane of polarisation, or perpendicular to it. For one position of the analyser one half of the field will appear dark and for another position the other half will be dark, and the angle between these positions is $2\alpha$, while for a position midway between them the two halves are equally illuminated. In this position the primitive plane of polarisation is parallel to PQ (Fig. 210). The experiment consequently consists in equalising the intensities by rotating the analyser, and as the position of equal intensity can be detected by the eye with much accuracy, the instrument affords an exact determination of the position of the primitive plane of polarisation.

Jellett employed this analyser in his saccharimeter to determine the rotation of the plane of polarisation produced by saccharine solutions, and he estimated[1] the probable error in the result of a single observation with such an instrument to be only 0·02 of a grain in a cubic inch of the solution.

**275. Laurent's Analyser and Saccharimeter.**—In Laurent's analyser[2] one half of the field of view is occupied by a plate of quartz, or gypsum, cut parallel to the axis, and of such a thickness that it introduces a retardation of $\frac{1}{2}\lambda$ between the ordinary and extraordinary rays. The other half of the field is empty or covered by a glass plate of sufficient thickness to absorb the waves of length $\lambda$ to the same extent as the crystalline plate.

Let APB (Fig. 211) be a semicircular plate of glass, AQB the plate of quartz, and AB the direction of its optic axis.

Then if the plane of polarisation of the incident light be parallel to

[1] Lloyd's *Wave Theory of Light*, p. 247. [2] *Journal de Physique*, 1874 and 1879.

OP, it will be transmitted without alteration through the glass and emerge from it still plane-polarised parallel to OP; but in the crystalline plate it will be divided into two components, one parallel to OA and the other perpendicular to it, and one of these will be retarded relatively to the other by a half-wave length. The result is that the light emerges plane-polarised from the crystalline plate, but the plane of polarisation will be inclined to OA on the other side at an angle AOQ = AOP (Art. 47, Cor. 2). The planes of polarisation of the rays emerging from the two halves of the compound plate will therefore be inclined at an angle $2\alpha$, where $\alpha = $ AOP. Consequently if a Nicol be fixed before the plate the two halves of the field will in general appear unequally illuminated, but if the principal plane of the Nicol be parallel to AB the illuminations of the two halves will be the same.

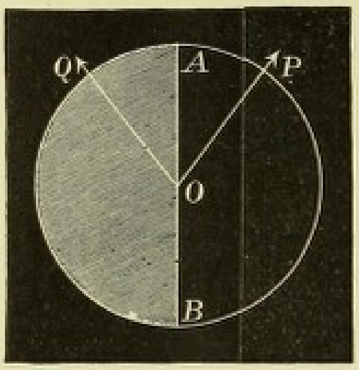



Fig. 211.

Since the thickness of a plate required to produce a retardation $\frac{1}{2}\lambda$ depends on the wave length, it follows that Laurent's analyser can be used only with monochromatic light. Jellett's analyser, on the other hand, is not subject to this restriction, but may be used with white light.

In Laurent's saccharimeter the light is sifted through a plate of bichromate of potash, which allows only the yellow rays to pass. It then falls upon the polarising Nicol, to the second face of which the quartz plate is fixed. After transmission through the plate it passes through a tube containing the solution under investigation, and thence through the analysing Nicol to an eyepiece where it is received by the observer.

**276. The Biquartz.**—A simple and fairly accurate means of determining the plane of polarisation is afforded by the biquartz. This consists of two semicircular plates of quartz placed in juxtaposition, each being cut at right angles to the axis. In one the rotation is right-handed, and in the other it is left-handed. The two plates are of the same thickness, so that they produce equal rotations of a given ray, but in opposite directions. Thus if QOP (Fig. 212) be the direction of the incident vibration, it will be rotated to the right in the plate APB, and by an equal amount to the left in the plate AQB. Now if the incident light be white the various simple colours will be rotated through different angles, and it follows that if the emergent light be analysed

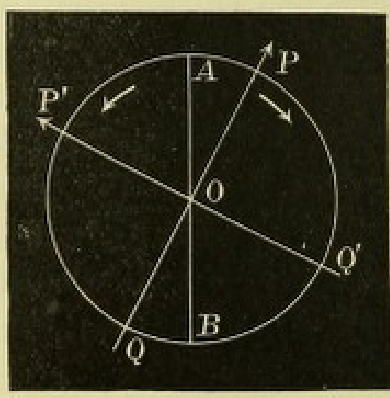



Fig. 212.

by a Nicol, waves of different refrangibilities will be quenched in the two halves of the field, and they will consequently appear of different colours.

There is one position of the Nicol, however, in which the two halves of the field may have the same colour. For consider the ray of which the plane of polarisation has been turned through a right angle. In one plate it will be turned to the right, and the vibration of this ray will be parallel to $OQ'$. In the other plate it is turned through 90° to the left, and the vibration will be parallel to $OP'$. Hence the vibrations of this ray will be parallel to the same direction in both halves of the field, viz. to a line $P'OQ'$ perpendicular to $POQ$. Thus if the principal plane of the Nicol be parallel to $POQ$ the light of the same wave length will be quenched in both halves of the field, and the other colours will be present in the same proportion in each half, so that they will exhibit the same tint.

By suitably adjusting the thickness of the quartz plates the ray which is rotated through a right angle may be made any one we please. If it be that corresponding to the brightest part of the spectrum (the yellow) this light will be absent and the common colour of both halves will be the tint of passage (see Art. 259). A minute rotation of the Nicol in one direction will render one half of the field blue and the other half red, and a displacement in the opposite direction will render the former half red and the latter blue. The thickness of the biquartz which exhibits the tint of passage is about 3·75 mm. When this is secured the plane of polarisation of the incident light is perpendicular to the principal plane of the analyser.

When the tint of passage is attained the substance under investigation is placed either between the polariser and the biquartz, or between the biquartz and the analyser. If the substance rotates the plane of polarisation the tint of passage will disappear from the field of the analyser, and to restore it the Nicol must be rotated in the direction in which the plane has been rotated and through an equal angle. In practice the light should fall normally on the biquartz, and the axis of rotation of the analysing Nicol should be parallel to the direction of the incident light. The greater part of the errors arising from these conditions not being accurately fulfilled may be corrected by rotating the Nicol through 180° and taking a second reading.

A method of greater accuracy is to subject the light to spectroscopic analysis. For this purpose it is thrown on the slit of a spectroscope by a lens so adjusted as to form a real image of the biquartz on the slit. Two spectra are seen in the field, one from each half of the biquartz and situated one above the other. Each of these is crossed by a dark band corresponding to that colour extinguished by the analysing Nicol. By rotating the Nicol the bands move across the spectra in opposite directions, so that the Nicol may be so adjusted that they are one above the other. When this is attained the light of the same wave length has been quenched in each half of the field, and the principal plane of the Nicol is parallel to the direction of vibration of the incident light. When the spectroscope is employed a single quartz plate may be used instead of the biquartz. For when solar light is used, and a spectrum formed which exhibits the Fraunhofer lines, the dark band arising from the quartz plate may be brought to coincide with any one of the dark lines of the spectrum. On introducing a rotatory substance the dark band will be displaced across the field, and to restore it to its original position the analysing Nicol must be turned through an angle equal to the rotation produced by the substance.[1]

**277. Poynting's Analyser.** — The leading idea in the design of analysers for the purposes of saccharimetry (as already illustrated in the foregoing examples) is to construct a piece of apparatus such that when a beam of plane-polarised light is transmitted through it, the light transmitted in one half of the field shall be polarised in a plane which is inclined at an angle to the plane of polarisation of the light transmitted through the other half. This, in general, is obtained by covering the two halves of the field with two pieces possessing rotatory power such that the plane of polarisation is rotated through different angles in passing through them.

The most obvious method of effecting this is to cover the two halves of the field with layers of different thickness of some rotating substance, for example, with two plates of quartz of different thickness cut perpendicularly to the axis, or with strata of a rotating liquid. This is the method employed by Professor Poynting.[2] In the first form of apparatus a uniform circular plate of quartz, cut at right angles to the axis, was divided along a diameter into two halves. One half was slightly reduced in thickness, and the two were then reunited so as to form a circular plate as before, with this difference, that one half of the plate now exceeded the other in thickness. When a beam of plane-polarised light is transmitted through such a plate the plane of polarisation of the light issuing from one half will be inclined to that of the light issuing from the other by an angle depending on the difference of thickness of the two halves of the plate.[3]

[1] M. Broch, *Ann. de Chimie et de Phys.* (3), tom. xxxiv. p. 119, 1852.

[2] J. H. Poynting, *Phil. Mag.* vol. x. p. 18, 1880.

[3] In order to vary this angle Professor Poynting suggests that one half of the plate might be made up of two wedges, as in Babinet's compensator (Art. 244), so that its thickness could be altered at pleasure by sliding one of the wedges over the other.

In the second form of apparatus a rotating liquid is employed, and the arrangement is exceedingly simple. The liquid, say a solution of sugar, is contained in a cell through which the beam of polarised light is transmitted. A piece of plate glass several mm. thick is immersed in this cell with its faces at right angles to the beam of light, and it is so arranged that one half of the beam passes through the plate while the other half passes by its edge. The former half of the beam, since it passes through the glass, obviously passes through a less thickness of the rotating liquid than the latter, and is consequently less rotated. Hence the light emerging from the two halves of the field is polarised in planes which are inclined to each other. The angle between the planes of polarisation can be varied by varying the thickness of the interposed glass plate.

**278. Soleil's Compensator and Saccharimeter.**—In the saccharimeter of M. Soleil the rotation produced by any substance under examination is compensated by means of a plate of quartz of which the thickness can be varied at will. This part of the apparatus is

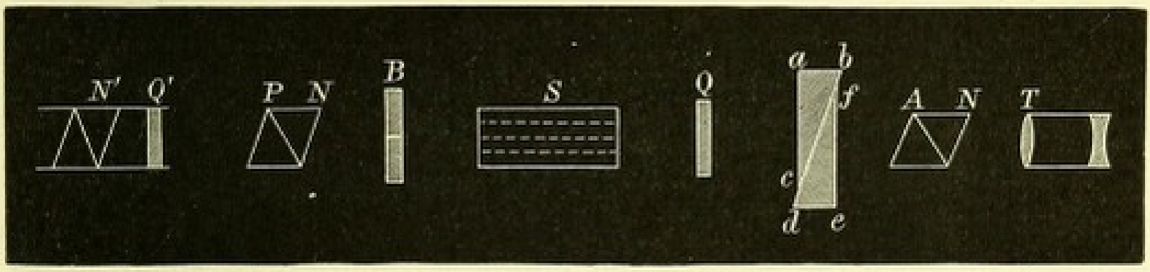



Fig. 213.

called the *compensator*, and is similar to Babinet's compensator (Art. 244) only in so far as it consists of two quartz wedges *abc* and *def* (shown exaggerated in Fig. 213), of which one can slide on the other along their common face so that their joint thickness may be varied. In this compensator, however, both wedges are cut with their faces *ac* and *ef* perpendicular to the axis, and they are both left-handed crystals. The light traverses both wedges in the direction of the optic axis, and the action of the compensator is consequently to introduce any desired amount of left-handed rotation. In conjunction with this pair of left-handed wedges a right-handed quartz plate Q is used, so that when the thickness of *abcd* is varied, the combined effect of Q and *abcd* is to introduce any desired amount of rotation, which may be either right-handed or left-handed according to the thickness of *abcd*. Thus when the thickness of *abcd* is properly adjusted, the right-handed rotation produced by Q will be exactly neutralised. For a greater thickness of *abcd* the combined effect of Q and *abcd* will be a left-handed rotation, while for less thicknesses the resultant effect will be right-handed.

In the saccharimeter the light is polarised by a Nicol's prism PN. It then falls upon a biquartz B, and afterwards traverses the substance S under examination. After emerging from the active substance it traverses a right-handed quartz plate Q and then falls upon the compensator *abcde*. Finally the apparatus is terminated by the analysing Nicol AN and a small Galilean telescope T, focussed on the biquartz B.

In making an experiment the tube S is filled with water,[1] and the telescope is focussed on the biquartz. The micrometer screw by which one of the wedges of the compensator is moved is now turned to the zero position, and the analysing Nicol is rotated till the tint of passage is obtained in the field. The water is now removed from the tube S and the rotatory substance is introduced. If the substance is active the plane of polarisation is rotated and the tint of passage disappears. It is, however, restored by turning the micrometer screw of the compensator so as to alter its thickness in such a manner that the rotation produced by the substance S is neutralised. If the substance is right-handed the thickness of the compensator must be increased, and if left-handed, diminished. The angle through which the micrometer has been turned in order to restore the tint of passage determines the amount of rotation produced by the substance.

This method requires the light to be white and the liquid colourless in order to obtain the tint of passage. To apply the instrument to cases in which the light or the liquid is coloured, M. Soleil added an extra piece which he called the *producer of the tint of passage*. This part is placed in front of the saccharimeter, and consists of a Nicol N′ and a quartz plate Q′. The incident light is polarised by the Nicol N′, and after traversing the rest of the apparatus, the field of the analyser will in general be coloured. By turning the Nicol N′ a position can generally be found in which the plate Q′ gives a tint complementary to the colour of the liquid for the light employed. The colour effect is thus neutralised and the tint of passage can be obtained.

**279. De Senarmont's Polariscope.**—In Soleil's compensator, just described, the characteristic piece consists of two wedges of left-handed quartz placed so as to form a plate of variable thickness. If, however, the wedges are of opposite sign so that one of them is right-handed while the other is left-handed, the plate will be such that along the central line the plane of polarisation will be rotated as much to the right in one prism as to the left in the other, and the resultant rotation along this line will be zero. Hence when the Nicols are crossed

[1] This avoids the necessity of readjusting the focus of the telescope when the rotatory liquid is introduced, for if the telescope were focussed on the biquartz when the tube S is empty, it would not be in focus when the tube is filled with a liquid.

this line will be black. On one side of this line the resultant rotation will be to the right, and on the other side it will be to the left, according to sign of the wedge of predominating thickness at the place in question, so that on either side of the central dark band coloured bands are symmetrically situated analogous to the bands observed in Babinet's compensator. When the analysing Nicol is turned the central band is displaced to one side or the other by an amount depending on the rotation of the Nicol.

This apparatus is known as De Senarmont's[1] polariscope, and it affords a very accurate means of determining the position of the plane of polarisation of any source of light which is partially or wholly polarised.

**280. MacCullagh's Theory of the Double Refraction of Quartz.**—The singular laws of the double refraction of quartz, and its peculiar property of rotating the plane of polarisation, were first connected in a general theory by MacCullagh.[2]

In a common uniaxal crystal two waves are propagated in any direction, one according to the ordinary laws of refraction, and the other in a manner depending on the inclination of the wave normal to the optic axis. These waves are polarised at right angles and travel with different velocities. Consider a plane wave passing through a crystal in a direction making an angle $\theta$ with the optic axis. Take the wave normal for the axis of $z$ and the axes of $x$ and $y$ in the wave front, the former being perpendicular to the optic axis. The axes of $x$ and $y$ will be parallel to the directions of vibration in the ordinary and extraordinary waves respectively (according to Fresnel's hypothesis), and if $\xi$ and $\eta$ be the displacements the equations of the two vibrations are

$$\frac{d^2\xi}{dt^2}=\mathrm{A}\frac{d^2\xi}{dz^2}, \quad \frac{d^2\eta}{dt^2}=\mathrm{B}\frac{d^2\eta}{dz^2} \qquad (1),$$

where, in a positive crystal ($a>b$), we have

$$\mathrm{A}=a^2, \quad \text{and } \mathrm{B}=a^2\cos^2\theta+b^2\sin^2\theta \qquad (2).$$

To account for the rotatory polarisation, and represent the twisted structure of quartz (p. 433), MacCullagh introduced differential coefficients of the third order into the equation of the vibrations, and assumed

$$\frac{d^2\xi}{dt^2}=\mathrm{A}\frac{d^2\xi}{dz^2}+\mathrm{C}\frac{d^3\eta}{dz^3} \qquad (3),$$

$$\frac{d^2\eta}{dt^2}=\mathrm{B}\frac{d^2\eta}{dz^2}-\mathrm{C}\frac{d^3\xi}{dz^3} \qquad (4).$$

[1] De Senarmont, *Ann. de Chim.* (3), tom. xxviii. p. 279, 1850.
[2] MacCullagh, *Trans. Roy. Irish Academy*, vol. xvii. 1836.

These equations are satisfied by

$$\xi = p\cos\frac{2\pi}{l}(vt - z), \quad \eta = q\sin\frac{2\pi}{l}(vt - z),$$

if, writing $q/p = k$, we have the equalities

$$v^2 = A - 2\pi Ck/l,$$
$$v^2 = B - 2\pi C/kl,$$

that is, if

$$A - B - \frac{2\pi C}{l}\left(k - \frac{1}{k}\right) = 0,$$

which by the formulæ (2) becomes

$$k^2 - \frac{l\sin^2\theta}{2\pi C}(a^2 - b^2)k - 1 = 0 —$$

a quadratic in $k$, of which one root is the negative reciprocal of the other.

Denoting the roots by $k$ and $-1/k$, we see that there are two elliptic vibrations as derived by Airy. Thus we have

$$\xi_1 = p\cos\frac{2\pi}{l}(v_1 t - z), \quad \eta_1 = pk\sin\frac{2\pi}{l}(v_1 t - z),$$

and

$$v_1^2 = A - 2\pi Ck/l.$$

This is an elliptic vibration propagated with velocity $v_1$ and having its axes in the ratio $1 : k$.

The second value $(-1/k)$ gives the vibration

$$\xi_2 = p\cos\frac{2\pi}{l}(v_2 t - z), \quad \eta_2 = -\frac{p}{k}\sin\frac{2\pi}{l}(v_2 t - z),$$

where

$$v_2^2 = A + 2\pi C/kl.$$

The axes of this vibration are in the ratio $-k : 1$, and it is propagated with a velocity $v_2$. Consequently we have in general two elliptic vibrations travelling with different velocities and having the same eccentricity, but the major axis of one coincides with the minor axis of the other.

When the light passes in the direction of the axis of the quartz we have $k = \pm 1$, which shows that there are two rays circularly polarised in opposite senses. Their velocities of propagation are

$$v_1^2 = a^2 - 2\pi C/l, \quad v_2^2 = a^2 + 2\pi C/l,$$

or approximately

$$v_1 = a\left(1 - \frac{\pi C}{a^2 l}\right), \quad v_2 = a\left(1 + \frac{\pi C}{a^2 l}\right).$$

The ordinary velocity $a$ is consequently their arithmetic mean.

Let the quartz be a parallel plate of thickness $e$. The time retardation of one ray relatively to the other in passing through the plate will be

$$\tau=\frac{e}{v_1}-\frac{e}{v_2}=\frac{e}{a}\left(1-\frac{\pi C}{a^2 l}\right)^{-1}-\frac{e}{a}\left(1+\frac{\pi C}{a^2 l}\right)^{-1}=\frac{2\pi e C}{a^3 l}.$$

Taking the velocity in air as unity, and the wave length $\lambda$, corresponding to $l$ in the crystal, we have $l:\lambda=a:1$, and the phase difference on emergence will be

$$\delta=\frac{2\pi\tau}{\lambda}=\frac{4\pi^2 e C}{a^4\lambda^2}.$$

But the rotation is half the phase retardation, therefore

$$\rho=\frac{2\pi^2 e C}{a^4\lambda^2},$$

which contains the experimental law of M. Biot, viz. that the rotation is inversely proportional to the square of the wave length. The sign of C determines whether the rotation is right-handed or left-handed, and the divergence from Biot's law exhibited by some substances might be accounted for by supposing that in these substances C is a function of the wave length.

## *Examples*

1. If the assumed equations of motion be

$$\frac{d^2\xi}{dt^2}=A\frac{d^2\xi}{dz^2}+C\frac{d^3\eta}{dz^3},$$

$$\frac{d^2\eta}{dt^2}=B\frac{d^2\eta}{dz^2}-C'\frac{d^3\xi}{dz^3},$$

show that $k$ is determined by the quadratic

$$k^2-\frac{l\sin^2\theta}{2\pi C}(a^2-b^2)k-\frac{C'}{C}=0.$$

This equation indicates that the ratio of the axes is not the same for both the elliptic vibrations, and explains the difference between the ratios observed by Airy.

2. Given

$$a=0{\cdot}64859,\quad b=0{\cdot}64481,\quad \lambda={\cdot}0000258 \text{ in.},$$
$$e=0{\cdot}03937 \text{ in.},\quad \rho=0{\cdot}333,$$

show that the quadratic for $k$ is

$$k^2-258\,k\sin^2\theta-1=0.$$

## CHAPTER XVIII

### ABSORPTION AND DISPERSION

**281. Selective Absorption.**—When light is being transmitted through any substance part of it is generally taken up by the medium or absorbed, and this absorption is different in general for rays of different refrangibility. It is on this account that white light generally becomes coloured after traversing a sufficient thickness of any substance. There is no body perfectly transparent—that is, transparent to rays of every refrangibility, and, on the other hand, there is no substance, in the same sense, perfectly opaque. Air[1] in sufficient thickness colours the light of the sun, at first yellow and afterwards red, and pure water produces the same effect in a more decided manner, while gold in a thin leaf transmits a faint greenish light. Metallic silver appears to allow the actinic rays to pass, and an ordinary stone wall is quite transparent to the long waves from a Hertzian vibrator (chap. xxi.).

The character of the absorption which takes place in any substance may be analysed by submitting the transmitted light to examination in a spectroscope. The spectrum formed by the transmitted light will be crossed by dark bands corresponding to the colours which have been most rapidly absorbed. Some substances exercise marked absorption in one or more parts of the visible spectrum, while others transmit freely the luminous rays, but absorb certain of the non-luminous waves. Thus the spectrum of light transmitted through a solution of chlorophyll (the colouring matter of the leaves of plants) exhibits dark bands in the red, yellow, green, and violet, and human blood produces marked absorption bands in the yellow and green portions of the spectrum, which are visible even when the quantity of blood is so small as to scarcely affect the colour of the solution. Permanganate of potash (Condy's fluid) exhibits several dark bands in the green part

Absorption bands.

[1] The transmitted light is the residue which passes through when absorption proper, and internal reflection arising from irregular structure of small suspended particles, have played their part. The relation of absorption to internal reflection deserves consideration.

of the spectrum, when the solution is dilute, and these are characteristic of the permanganates. With stronger solutions the dark bands broaden out, and ultimately unite to form one broad dark band, the whole region being now absorbed. And again, Brewster discovered that a solution of oxalate of chromium and potash produces a single narrow absorption band, which is so well defined that he suggested its use in the measurement of refractive indices with artificial light, when sunlight cannot be obtained.

For the most part, however, the light which has passed through an absorbing medium exhibits a single maximum of absorption. The rays in a certain region of the spectrum are most effectively stopped; but in the case of gases the phenomena are very different. The rays stopped by them are characteristic and well defined, so that the spectrum of the transmitted light exhibits a certain number of fine dark lines, distinctly separated from each other. Gases accordingly allow almost all the light to pass except a few rays of definite refrangibility. Gases.

When light of any wave length is intercepted by an absorbing substance the energy of the corresponding ether vibrations is imparted to the matter, and in general reappears as heat, as is manifested by the rise in temperature of the absorbing substance. The vibrations of the ether are taken up by the matter molecules, and these in turn vibrate and become centres of disturbance. What is absorbed then is energy. The matter molecules are intimately bound up with the ether, and the energy of the ether vibration is converted (possibly) into energy of motion in the matter molecules.

Now if we consider any system of matter molecules it will be easily inferred that some ether vibrations will be much more powerfully absorbed than others, viz. those most competent to excite that vibration which the matter molecules execute freely. It is well known that a series of slight taps may excite a considerable oscillation in a common pendulum, the condition being that the taps be timed to the period of swing of the pendulum. The same point is illustrated in the resonance of organ pipes which are excited by vibrations of their own free period. So, again, if a number of ships at sea be disturbed by waves, each ship will have its own free period, and if for some of them this happens to be the same as that of the waves, these particular ships will be thrown into a state of violent oscillation. They will thus act as absorbers of the energy of the waves. If the periods are not synchronous the motion caused by one wave will be destroyed by another, so that the energy absorbed from one wave will be given out to some other, and there will be no accumulation of energy in the ship. Analogy.

It consequently follows that of the multitude of waves in a pencil of white light those which will be most freely absorbed by any substance will be those which synchronise in period with the free vibrations of the matter molecules. The molecules absorb waves of their own period of vibration, and these again are obviously the waves which the molecules will excite when they become centres of disturbance.[1] It is consequently to be expected that different kinds of matter should absorb waves of different periods or rays of different refrangibility. That they should, in fact, select certain rays in preference to others, and this is what is known as *selective absorption*.

Selection.

In the case of sounding bodies we know that some can take up and resound to vibrations of almost any period, while others confine their vibrations to some fundamental tone and its harmonics. The former class is illustrated in the sounding board of a piano or the body of a violin, while examples of the latter are an organ pipe or a stretched string. So it is in the case of absorbing substances. Some, such as black opaque bodies, absorb the luminous waves of every length, and others, such as gases, absorb only waves of certain definite periods. When an opaque substance, like a piece of iron, is heated, at low temperatures it emits the longer or dark heat waves. As the temperature is raised the vibrations become more intense, the molecular agitation more excited, and waves of shorter period are emitted, so that at a certain stage the waves of the red end of the spectrum appear and the iron mass assumes a dull red heat. Carrying the elevation of temperature still further, vibrations of shorter and shorter periods are superadded, till finally all the waves of the visible spectrum are emitted and the mass has reached a *white* heat.

Gases.

In the case of a gas, on the other hand, the molecules have free paths ; they are at distances from each other which are large compared with the dimensions of the molecule, and each, in making its excursions between its successive impacts with the others, will vibrate in its own free period of oscillation. The waves absorbed by a gas will therefore be well defined and of certain definite periods, viz. those of the free vibration peculiar to the molecule. And these, again, will be of the same period as those emitted by the gas when it becomes incandescent —that is, the spectrum of an incandescent gas will consist of one or more well-defined bright lines, images of the slit, which correspond exactly to the dark absorption lines which appear in the spectrum of white light which has passed through the gas when cold. Thus the

[1] Tourmaline rapidly absorbs the ordinary wave—that is, it absorbs light which is plane-polarised in a certain direction, and when heated, so as to become in turn a source of light, it radiates plane-polarised light.

spectrum of incandescent sodium vapour contains a bright double yellow line, and the spectrum of white light which has passed through cold sodium vapour is crossed by a dark double line in the yellow corresponding exactly to the bright yellow line given by the vapour when hot. The spectrum of a given gas may, however, vary considerably with changes of temperature and pressure.

Thus hydrogen enclosed in a vacuum tube at atmospheric pressure exhibits, with a certain amount of continuous spectrum, four bright lines, three of them situated in the red, blue, and violet respectively, and the fourth, a faint one, in the extreme violet. As the pressure is decreased they grow sharper and more distinct, while the continuous part of the spectrum grows fainter and dies away. If, however, the exhaustion be pushed further, the red and violet lines fade out and the green alone is left. On the other hand, if more hydrogen be compressed into the tube, the bright lines will be found to broaden out and grow fainter, while the continuous spectrum becomes more brilliant. So, again, oxygen at high temperatures—that is, with a strong spark—shows a number of bright lines chiefly in the violet, but if the temperature be lowered only four lines are presented,—one in the red, two in the green, and one in the blue.

Increase of temperature in general introduces new lines, and increase of pressure causes them to grow wider and broaden out into a continuous spectrum. This is not contrary to expectation, for increase of temperature corresponds to increased molecular agitation, and increase of pressure brings the molecules closer together, so that collision becomes more frequent. During collision the vibrations are constrained and the emitted waves are altered. The denser the gas the greater is the proportion of time spent in collision and the less the free path of the molecule, so that the free or characteristic vibrations of the molecule grow less and less predominant in the emitted light. The result is that when the gas is compressed to near its liquid volume the irregular vibrations have gained the field, and the characteristic lines of the gas have broadened out into a continuous spectrum like that of a solid or liquid.

**282. Coefficients of Transmission and Absorption.**—When a beam of light of a given wave length is transmitted through a layer of an absorbing substance a certain fraction of it is absorbed, and it is assumed that this fraction is independent of the intensity of the incident beam. It follows from this that the amount of light which passes through a number of equal layers diminishes in geometrical progression as the number of layers increases in arithmetical progression. Thus if I denotes the intensity of the incident light, $I\alpha$

will be the intensity after transmission through unit thickness, where $\alpha$ is a proper fraction and depends upon the nature of the substance and the refrangibility of the light employed. For a given wave length $\alpha$ will be different for different substances, and for a given substance $\alpha$ will vary with the wave length. The intensity of the beam incident on the second unit of thickness being $I\alpha$, it follows that the intensity of the beam transmitted through it will be $I\alpha \cdot \alpha = I\alpha^2$. Similarly the intensity after transmission through three, four, etc., units will be $I\alpha^3$, $I\alpha^4$, etc., respectively, and the intensity after transmission through a thickness $x$ will be

$$I' = I\alpha^x.$$

The quantity $\alpha$ is termed the *coefficient of transmission*. If the incident pencil be not monochromatic, but be composed of waves of various lengths, of which the intensities are $I_1$, $I_2$, $I_3$, etc., with coefficients of transmission $\alpha_1$, $\alpha_2$, $\alpha_3$, etc., the intensity of the transmitted beam will be

$$I' = I_1\alpha_1^x + I_2\alpha_2^x + I_3\alpha_3^x + \ . \ . \ . = \Sigma I\alpha^x.$$

The quantity, and the quality, of the transmitted light will consequently depend upon the primitive composition of the beam—that is, upon the nature of the source; also upon the nature of the absorbing substance—that is, upon the various coefficients $\alpha_1$, $\alpha_2$, etc.; and finally upon the thickness traversed.

If, on the other hand, we deal with the absorption so that a beam of intensity I is changed by an amount $-dI$, when transmitted through a layer of thickness $dx$, we assume that $dI$ for light of a given wave length is proportional to $dx$, and also to I, so that we have

$$-dI = \beta I dx,$$

where $\beta$ is a constant depending on the nature of the substance, and on the wave length of the light. This constant is termed the *coefficient of absorption* of the substance for the light in question. Integrating this equation we have at once

$$\log. (I/I_0) = -\beta x,$$

where $I_0$ is the intensity of the incident beam, and I the intensity after it has traversed a thickness $x$ of the substance. Writing this equation in the form

$$I = I_0 e^{-\beta x},$$

we see that the coefficients of absorption and transmission are connected by the relation $\alpha = e^{-\beta}$.

**283. Dichromatism—Change of Tint—Critical Thickness.**—The coefficients of transmission being in general different for the different colours, it follows that the emergent light will be coloured, the colour being generally the same for all thicknesses traversed, and the tint merely becoming deeper as the thickness is increased. There are, however, some media in which the colour of the transmitted light varies with the thickness. Thus if $I_1$ and $I_2$ be the initial intensities of the two colours which predominate in the transmitted light, $\alpha_1$ and $\alpha_2$ their coefficients of transmission, then their intensities in the transmitted beam will be $I_1\alpha_1^x$ and $I_2\alpha_2^x$. Consequently, if we have $I_1>I_2$ and $\alpha_2>\alpha_1$, it will follow that for small thicknesses $I_1\alpha_1^x$ will be greater than $I_2\alpha_2^x$, but as the thickness increases these quantities will become more nearly equal, till at last exact equality will be obtained, and for greater thicknesses $I_2\alpha_2^x$ will be greater than $I_1\alpha_1^x$. The thickness at which equality is reached is determined by

$$I_1\alpha_1^x = I_2\alpha_2^x,$$

or, taking logarithms of both sides, we have

$$x = \frac{\log. I_2 - \log. I_1}{\log. \alpha_1 - \log. \alpha_2},$$

which determines the thickness at which the change of colour takes place. For thickness less than this $I_1$ predominates, but for greater thicknesses $I_2$ predominates. For example, cobalt glass transmits both blue and red light, but the blue to a less extent than the red. For this reason, when the thickness is considerable, the blue rays of the spectrum are almost entirely absorbed, and the colour of the transmitted light is red. The glass appears blue on the other hand when the thickness is small.

**284. Colour produced by Absorption.**—From the foregoing considerations it is clear that absorption must be a prime agent in the production of colours in natural bodies. When white light passes through any substance some of its components are wholly or partially absorbed, and the transmitted beam is more or less coloured in consequence. The tint depends on the nature of the missing rays, and is the resultant of the rays which have been allowed to pass. Ordinarily bodies are seen by light which has been scattered at their surfaces by rugosities or inequalities. Structural inequalities in the interior also scatter the light which has penetrated beneath the surface, so that it is reflected back and reaches the eye, after having traversed some thickness of the material, weakened and robbed of some of its constituents by absorption.

An excellent illustration of the effect of interior scattering is presented in the striking dissimilarity in appearance between a liquid and its froth. Pure colourless water gives a white foam, clear glass and ice may be powdered into white dust, and the white froth on beer scarcely exhibits a trace of the deep colour possessed by the liquid.

In all these cases the incident light suffers partial reflection at the surface of each bubble or particle on which it falls. The intromitted rays are almost all returned by partial reflections at a multitude of surfaces, and emerge again from the substance in all directions. The consequence is that such accumulations of particles or films are highly opaque to light, and if they exercise no particular absorbing effect on any of the simple colours they will reflect all colours equally and appear intensely white; such, for example, is a cloud. If, however, the material absorbs some of the simple colours more powerfully than others the emergent light will be coloured, and this we designate the colour of the body. It may, however, happen that the substance absorbs some of the rays more readily than others, and yet in a finely divided state it appears white because nearly all the light is reflected very close to the surface, and does not pass through a sufficient thickness of the material to sensibly tinge the scattered light. This happens in the case of the froth of beer just mentioned, or when blue or red glass is finely pounded, the resultant powder being nearly white. Some substances, however, absorb particular rays so strongly that passage through a very thin film may tinge the scattered light in a decided manner.

It is interesting to note the variations in appearance produced when a coloured body is placed successively in the different parts of a pure spectrum. A white lily placed in the red looks red, in the blue it appears blue, and placed in the green it exhibits a green colour. The white lily contains no fluid which absorbs any particular colour more readily than any other. It affects all wave lengths alike so that it assumes the colour of the light that falls upon it, and in white light it appears white.

A coloured flower, on the other hand, when held in the different parts of the spectrum, will appear of the same colour as that part in which it is held, but it will vary much in brightness as it is moved from one part to another. Placed in a colour which it does not absorb it will appear bright, for all the light that falls upon it will be reflected and traverse the flower cells without sensible diminution, but in a colour which it absorbs freely it will appear almost black, for nearly all the incident light will be absorbed. Thus a red poppy will show a brilliant red when placed in the red part of the spectrum, but it will appear dark in the green or blue. Its cells are filled with a fluid which absorbs the green and blue, but which freely transmits the red.

The effect of internal reflection, and subsequent absorption, in producing colour is well shown by placing a carefully filtered coloured liquid in a basin. The light falling on the surface of the liquid penetrates the interior, and after reflection at the sides of the basin it emerges and reaches the eye, having suffered absorption, and the coloured liquid is seen. If, however, the interior of the basin be painted black, no light will be reflected at its sides, and consequently none will reach the eye from the interior, and the liquid will appear black. The upper surface of the liquid reflects all the rays in the same proportion. Any selective reflection of the surface could be detected by observing in it the image of a white object. This will appear coloured if the rays are reflected in different proportions. By sprinkling a little flour or powdered chalk in the solution, the full colour may now be restored, for the light which enters the liquid will be reflected from the white particles, and, emerging, will reach the eye, robbed of those rays which are absorbed most copiously by the solution.

Reflection then is the proximate cause of the colour of these bodies, inasmuch as without reflection no light would reach the eye, but absorption is the ultimate cause, for it is thus that the reflected light is deprived of some of its constituents and becomes coloured.

At this point it will be easy to understand the colour displayed by a mixture of pigments. From what has been said the mixture should present a tint resulting from the combined effect of the colours transmitted by all the pigments. Thus a mixture of blue and yellow paints appears green, not because green is situated between blue and yellow in the spectrum, but because green is the only colour which is freely transmitted by both. The blue pigment absorbs the red, orange, and yellow rays, while the yellow paint absorbs the violet, indigo, and blue. Hence the green alone is suffered to pass through both and emerge to the eye after internal reflection.

Many bodies, however, possess *surface colour*, and seem to select some rays for reflection in preference to others. Thus many of the aniline dyes appear one colour when viewed by reflected light and another when viewed by transmitted light. To observe this a small quantity of a solution in alcohol may be spread on a glass plate and allowed to evaporate. A thin film of the aniline is thus deposited on the plate, and this will present one colour when looked at and another when looked through. A similar selection appears to take place in reflection from metals, gold reflecting yellow and copper red. A large number of substances have been found to possess this surface colour, appearing to refuse certain rays admission altogether.

The light transmitted through such a substance will consequently be deficient from two causes. It will be deprived of certain rays by reflection at the surface and of certain others by absorption in the interior, and the constitution of this doubly sifted beam will determine the colour by transmission, or the body colour of the substance.

On the other hand, the light reaching the eye by reflection consists of two parts, one the abnormal portion which has been refused admission and the other scattered in the ordinary manner after penetrating the substance a little below the surface and suffering absorption. This mixture determines the colour as seen by reflection, and when examined through a Nicol's prism shows a marked change in character. The ordinary reflected part may be plane-polarised at a proper angle of incidence, and can be extinguished by the Nicol, but the anomalous part of the beam, as in the case of metallic reflection, and for similar reasons, is never plane-polarised, and will therefore be always visible through the Nicol. The result is that the colour of the surface will change as the analyser is turned. Thus fuchsine appears a rose colour by transmitted and green by reflected light, but when examined through a Nicol it appears a peacock blue.

**285. Absorption Lines in the Solar Spectrum.**—When a pure spectrum of the sun's light is obtained it is found to be crossed by an enormous number of very fine dark lines. These lines were first noticed by Dr. Wollaston, but they were subsequently rediscovered and investigated with great skill by Fraunhofer, after whom they are known as "Fraunhofer's lines." It is easy to understand how this want of continuity in solar light escaped notice till great precautions were taken to secure a comparatively pure spectrum. For if the waves emitted by the sun were all of one period—that is, if its light were perfectly homogeneous—then on transmission through a prism there would be no separation or dispersion, and in the spectroscope a narrow image of the slit would be seen; while if the light contained only waves of two definite lengths, say red and blue, then two narrow images of the slit would be seen, one red and the other blue, and these would be separated by a dark space, the width of which will depend upon the dispersive power of the prism or the resolving power of the spectroscope.

With light composed of many colours we will have as many coloured images of the slit arranged parallel to each other, and separated by dark spaces corresponding to the waves which are not represented in the original beam. When the constituent waves of the composite beam become sufficiently numerous the various images of the slit will be in close proximity, and a stage will be arrived at when they will unite to form a continuous band or even overlap each other. This stage may obviously be postponed by increasing the dispersion—that is, by employing several prisms—and also by diminishing the width of the slit.

The number of different waves in a composite beam may, however, be so great as to defy resolution by any apparatus which we can construct, and the spectrum will then appear quite continuous; but we are not, therefore, to conclude that it comprises light of all possible wave lengths, but merely that its constituents are too numerous to be separated by our apparatus. Consequently, whether a spectrum appears continuous or not will ultimately depend upon the resolving power of the instrument employed.

The existence of dark lines crossing the solar spectrum shows us that the corresponding waves are either wholly or largely absent from the light of the sun when it reaches us, and the inference is that they have either been absorbed by some medium between us and the sun, or else they have not been emitted at all.

Now the spectra of incandescent solids, such as iron, charcoal, etc., do not exhibit the numerous dark lines seen in the solar spectrum, and if the sun be an incandescent mass we infer that he emits the missing rays, that they are intercepted by some absorbing medium, and that this medium is the atmosphere of cooler gases surrounding the inner nucleus of the sun.[1] Thus the dark D lines in the yellow part of the solar spectrum indicate that the light from the inner part of the sun passes through a stratum of sodium vapour in the outer atmosphere, for it is found that these lines correspond exactly to the absorption bands of sodium vapour, and that the corresponding rays are emitted by the vapour when incandescent.

By comparing the other dark lines with the absorption lines produced in continuous spectra by other vapours, or with the bright lines afforded by them when incandescent, the existence of many other substances has been detected in the sun, and the examination of the constitution of the heavenly bodies has been rendered possible. The spectrum of iron vapour consists of a large number of bright lines, and of these over four hundred have been identified with dark lines in the solar spectrum, and corresponding coincidences of the dark lines with the bright lines given by many

[1] If all the dark lines were the effect of absorption while passing through the earth's atmosphere or through some substance existing between the earth and the sun, then the spectra of all the stars should exhibit absorption bands at least as extensive as those of the solar spectrum. Lines produced by absorption in the earth's atmosphere should be common to all spectra, and should vary in intensity according to the thickness traversed by the light—that is, at sunrise and mid-day.

other substances have been observed. Thus Ångström and Thalén identified the following:[1]—

| | | | | | |
|---|---|---|---|---|---|
| Iron . . | 450 | Manganese . | 57 | Hydrogen . | 4 |
| Calcium . . | 75 | Chromium . | 18 | Aluminium . | 2 |
| Barium . . | 11 | Nickel . . | 33 | Zinc . . | 2 |
| Magnesium . | 4 | Cobalt . | 19 | Copper . . | 7 |

Besides the examination of the resultant light from the sun as a whole, that emitted by any selected portion of it may also be examined. An image of the sun an inch or two in diameter may be formed by means of a lens of considerable focal length, and the slit of the spectroscope may be adjusted to any desired part of the image, and the radiations from that part may be separately studied. This has been done by Janssen, Lockyer, and many others, with the result that a remarkable variation of appearance occurs, especially in the neighbourhood of spots and faculæ.[2] Sometimes some of the lines corresponding to an element are seen bright and straight, while others are faint or crooked.

**286. Width of the Spectral Lines—Radiation from Moving Molecules.**—In discussing the radiation from an incandescent gas, we have so far neglected the motions of translation of the individual molecules. We have supposed that the molecules emit waves of certain definite periods depending upon their own period of free vibration. In this case the spectrum of an incandescent gas should consist of certain definite lines, each of purely monochromatic light, and the absorption bands produced in a continuous spectrum should be of infinitesimal width. Taking into account the motions of translation of the molecules while describing their free paths, and applying Doppler's principle, it will follow immediately that the spectral lines exhibited by an incandescent gas as well as the absorption bands in the solar spectrum should be of definite width. This application has been discussed by Ebert,[3] and the whole investigation with reference to interference has been taken up and extended by Lord Rayleigh.[4]

[1] [For the results of Rowland's great work on the solar spectrum see "Preliminary Table of Solar Wave-Lengths," *Astrophysical Journal*, vol. i. p. 29 (1895), and vols. following.]

[2] [At the Exeter Meeting of the British Association (1869) Janssen described a method of studying the distribution of any substance in the sun which produces a brilliant line in the spectrum. A large image of the sun is formed on a movable slit and a spectrum is produced by a grating. A second slit, moving in correlation with the first, allows the bright line alone to act on a photographic plate immediately behind it. In this way Mr. George Hale, and more recently M. Deslandre, have been most successful in obtaining photographs of the sun produced solely by the K line (see Hale, *Astronomy and Astrophysics*, vol. xii. (1893), pp. 241, 450).]

[3] Ebert, *Wied. Ann.* xxxvi. p. 466, 1889.

[4] Lord Rayleigh, *Phil. Mag.* vol. xxvii. pp. 299, 484, 1889.

Thus if $v$ be the velocity of light and $u$ the velocity of the moving molecule, $\theta$ the angle which its direction of motion makes with the line of sight, then the natural wave frequency $n$ is changed into $n'$, where

$$n' = \frac{nv}{v - u\cos\theta}.$$

The new wave length $\lambda'$ will be $v/n'$, and will consequently be connected with the old wave length $\lambda = v/n$, by the equation

$$\lambda' = \lambda(1 - \frac{u}{v}\cos\theta)$$

where $u$ is small compared with $v$.

As a first approximation $u$ may be supposed to be the same for every molecule, and the limiting values of the frequencies will be

$$n\left(1 + \frac{u}{v}\right), \quad \text{and } n\left(1 - \frac{u}{v}\right).$$

The result being that the spectral line which would otherwise be infinitesimally narrow, and correspond to a single wave frequency $n$, will now be broadened out into a uniformly illuminated band, having the above limits for the frequencies at its edges. Thus by the motion of the molecules the emitted light is rendered to some extent heterogeneous. The waves are not of a single period, but vary over a certain range depending on the velocity of motion of the molecules, and conversely, the light absorbed by a gas will not be of a single definite wave length, but will consist of a group or groups of waves lying between certain limits.

## ANOMALOUS DISPERSION

**237. Abnormal Spectra.**—Intimately connected with the dark absorption bands is the remarkable phenomenon of anomalous or abnormal dispersion. Hitherto we have always spoken of the spectrum as presenting an invariable succession of colours in the order violet, blue, green, yellow, orange, and red. We have incidentally mentioned that spectra obtained by prisms of different substances present peculiarities characteristic of the material of the prism, and that spectra produced by different prisms cannot well be compared on account of the irrationality of dispersion, as the irregularity in the spreading out of the colours is termed.

It has been found that refraction spectra are not complicated by the irrationality of dispersion alone, but that in many cases the order of the colours is entirely changed, and that sometimes the spectrum may not even present a continuous appearance, but may be broken into parts isolated from each other by broad dark bands.

Although anomalous dispersion may appear at first sight to be a strange and unexpected phenomenon, yet, on duly considering the mode of propagation of light in refracting media, we are led to expect it. For the refractive index of any medium, for waves of a given frequency, is determined by the velocity of propagation, and in general the velocity of propagation will vary with the wave length in a manner depending on the nature of the medium. Thus if the long waves happen to travel faster than the short waves in some media, we are not justified in concluding that there are not media in which the reverse may be the case. The velocity of waves of a given frequency is influenced by the internal structure of the medium, and the constitution of one medium might be such that the violet waves travel faster than the red, while the reverse may be the case in another medium, just as in one medium the red may be more strongly absorbed than the violet, while in others the violet is more strongly absorbed than the red. The so-called irrationality of dispersion, and anomalous dispersion, are consequently phenomena which ought to be expected just as we should expect the different colours to be absorbed in different proportions by a given medium, and differently in different media.

The existence of *anomalous dispersion* seems to have been first discovered by Fox Talbot[1] about 1840, but the discovery does not seem to have been followed up. In 1860 M. Le Roux[2] discovered that iodine vapour possessed a very remarkable absorbing power. He found that it transmitted only the red and violet rays, and that of these the red are the more refracted, contrary to what takes place in the ordinary cases of refraction. At a temperature of 700° C. the indices for iodine vapour of the red and violet rays were found by M. Hurion[3] to be

$$\mu_r = 1{\cdot}0205, \quad \mu_v = 1{\cdot}019.$$

In 1870 Christiansen[4] detected the existence of anomalous dispersion in alcoholic solutions of fuchsine (one of the aniline dyes). This solution gives a marked absorption band in the green, so that this colour is entirely absent from the spectrum of the transmitted light. Of the remaining colours the red, orange, and yellow occur in their natural order—that is, the red is less refracted than the orange, and the orange less than the yellow. The violet, however, suffers a very peculiar deviation. It is

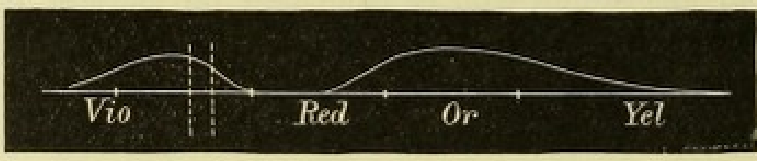


Fig. 214.

[1] Tait, *Light*, § 196, and *Proc. Roy. Soc. Edinburgh*, 1870-71.
[2] Le Roux, *Ann. de Chimie et de Physique*, third series, tom. lxi. p. 285, 1861.
[3] M. Hurion, *Journal de Physique*, first series, tom. vii. p. 181.
[4] Christiansen, *Pogg. Ann.* 1870-72.

less refracted than the red, and separated from it by a dark interval. The arrangement and relative intensities are shown in Fig. 214, the spectrum is elongated to an extraordinary extent, the green is absent, the violet is least refracted, and is separated from the other colours by a dark space. The prism used by Christiansen consisted of two glass plates inclined at an angle of about 1°. The dotted lines show the position and length of the spectrum obtained with the same prism when filled with pure alcohol.[1]

The indices for the various rays were as follows :—

| A | B | C | D | E | F | G | H |
|---|---|---|---|---|---|---|---|
| Fuchsine Solution (18 per cent) | 1·450 | 1·502 | 1·561 | ... | 1·312 | 1·258 | 1·312 |
| Pure Alcohol | 1·3628 | not measured | 1·3654 | 1·3675 | 1·3696 | 1·3733 | 1·3761 |

In determining the indices by the method of total reflection, it is found at normal incidence that the reflected light is strongly coloured green. The dark band in the green of the transmitted light is consequently accounted for. As the incidence is increased the spectrum of the reflected light shows first (besides the green) blue, then violet, red, orange, and yellow enter in succession. We conclude, therefore, that the green is totally reflected at all incidences, and that of the other colours the blue is the least refracted and the yellow most.

In working with a hollow prism containing a solution, the action of the solvent may be eliminated by enclosing the prism in a vessel having parallel glass sides and filled with the solvent liquid.

Kundt[2] has made an extensive series of observations on anomalous dispersion, and has shown that it appertains to all bodies which possess what is known as *surface colour*—that is, whose colour by reflection is different from the colour by transmission (p. 473). Thus fuchsine colours the transmitted light red, but the light reflected from it is green. All bodies of this class totally reflect certain colours of the spectrum at all incidences, and their solutions, even though very dilute, show marked absorption bands in the same colours. Among the substances examined by Kundt were the blue, violet, and green anilines, solution of indigo in strong sulphuric acid, carmine, permanganate of potash, and cyanine.

The greater part of the observations were made by the method of crossed prisms. This method has been employed by Newton in his

[1] [See also papers by A. Pflüger, who used prisms of the solids of very small angle 40″ to 140″, in *Wied. Ann.* vol. lvi. p. 412 ; vol. lviii. p. 670 ; vol. lxv. pp. 171, 214.]

[2] Kundt, *Pogg. Ann.* 1871-72.

investigations on the refrangibility of solar light (see chap. v. p. 120). If two prisms of a substance, such as glass, which exhibits regular dispersion, be crossed, the spectrum afforded by the first will be displaced more or less by the second, but the displacement will affect all the colours in a continuous manner, the violet most and the red least. If, however, one of the prisms exhibits anomalous dispersion, the displacements may take place in opposite directions, and be very different for the various colours. The displaced spectrum will then consist of coloured patches more or less isolated from each other. Fig. 215 shows a spectrum of permanganate of potash obtained by the method of crossed prisms. It exhibits five absorption bands in the green, and is inflected in contrary directions at the two extremities.[1]

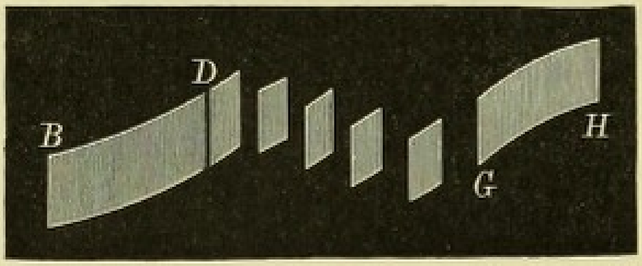


Fig. 215.

**288. Kundt's Law.**—As a result of his experiments M. Kundt concluded that in going up the spectrum, from the red towards the violet, the deviation is abnormally increased below an absorption band, while above the band the deviation is abnormally diminished by the absorption. Or in going up the spectrum the refractive index is increased where the coefficient of absorption increases rapidly, and diminished where the absorption rapidly diminishes.

Thus on either side of an absorption band there is an abnormal change of refrangibility of such a kind that the refraction is *increased below*, that is on the red side, and *diminished above* the band. An analogous interaction frequently occurs in vibrating systems of nearly the same period. Thus if to the bob of a pendulum P, executing

[1] [Professor H. Becquerel (*Comptes Rendus*, vol. cxxvii. p. 399, December 5, 1898; and vol. cxxviii. p. 145, January 16, 1899) has recently exhibited the anomalous dispersion of sodium vapour near the D lines. He projected the image of the crater of an arc light on a horizontal slit, placed in the focus of a collimator lens. The parallel beam then passed through a sodium flame in the form of a prism with its refracting edge horizontal, and was focussed by a lens into an image of the horizontal slit upon the vertical slit of a highly dispersive spectroscope. On either side of the two dark D lines the spectrum was sharply curved so that for wave-lengths slightly exceeding those of $D_1$ and $D_2$ the vapour possessed an index of refraction rapidly increasing near the lines; and for wave-lengths slightly less than $D_1$ or $D_2$ the index decreased when approaching the lines (see Lord Kelvin, *Phil. Mag.* March 1899, and W. H. Julius, *Proc. of the Royal Acad. of Sciences*, Amsterdam, or *Astrophysical Journal*, vol. xii. p. 185, October 1900). In his interesting paper Julius suggests that the extraordinary curvature of certain solar lines may be due to anomalous dispersion in portions of the sun's atmosphere.]

horizontal vibrations, another pendulum $p$ be attached, the effect will be to increase or diminish the period of P according as the period of $p$ is shorter or longer than that of P. That is, the effect of $p$ is to increase or diminish the virtual inertia of P according as the natural period of the former is shorter or longer than that of the latter. If P tends to vibrate more rapidly than $p$, then the effect of $p$ is to make it vibrate more rapidly still, and *vice versâ*. Below an absorption band the ether vibrations are slower than those of the matter molecules. The effect of the matter is consequently to increase (abnormally) the virtual inertia of the ether, and therefore the refrangibility. Above an absorption band the period of the matter is longer than that of the ether, and the effect is an abnormal diminution of the virtual inertia of the ether, and therefore of the refraction.[1]

## Transformation of Radiations

**289. Fluorescence.**—In discussing the phenomena of absorption it has been pointed out that of the many constituents of a beam of solar light some are absorbed more particularly by one body and some by another. In general the energy of an ether wave will be taken up by the matter if it is of the proper period to excite the matter molecules to vibration, yet some bodies appear competent to take up and absorb the energy of waves of all periods, at least of those lying within the limits of the visible spectrum. The general effect of absorption is to raise the temperature of the body. Thus light is taken up and heat is emitted instead—that is, waves of a short period are absorbed—and in their place the longer infra-red or heat rays are emitted. This is obviously a case of the transformation of radiations of a high period into others of a lower period. Energy is absorbed in the form of short waves and is emitted in long waves, and this species of degradation is always in operation when light falls upon and heats material bodies. We shall see immediately that light waves of a short period, as those from the violet end of the spectrum, are absorbed by some substances and emitted again as light waves of a lower refrangibility. For example, the violet, or even the ultra-violet, may be specially absorbed and emitted again as green or red rays. This is degradation of the same sort as the former, and is termed *fluorescence*. When light is absorbed by lamp-black and emitted as heat radiations, we detect the transformation by our sense of heat, but in the case of fluorescence the degradation is seen by the eye; the change in period has not been

[1] This principle was used by Lord Rayleigh as early as 1872. See *Phil. Mag.* 1872, "On the Reflection and Refraction of Light by intensely opaque Matter."

sufficiently great to place the vibration outside the limits of the visible spectrum.

**290. Calorescence.**—The converse transformation may also be effected, viz. the conversion of waves of long period into those of shorter period. The long heat waves may be converted into waves sufficiently short to affect the eye, and may thus be rendered visible. This converse transformation, the rendering visible of the infra-red waves, is termed *calorescence.* It may be demonstrated by concentrating the non-luminous radiations by means of a condensing lens at the focus of which a piece of platinum foil is placed. Professor Tyndall first exhibited it in this manner by focussing the radiation of an electric arc on a slip of platinised platinum. The beam was sifted of all the luminous radiations by transmission through a solution of iodine in bisulphide of carbon, so that only the obscure heat radiations fell upon the platinum. In this manner the slip was raised to incandescence and emitted light—that is, the visible infra-red radiations of long period were converted, in part at least, into the higher luminous radiations.

**291. Fluorescent Substances — Stokes's Theory.**—The phenomenon of fluorescence was first observed by Sir David Brewster[1] in an alcoholic solution of chlorophyll, and was termed by him internal dispersion. He found that when a pencil of solar light passes through a green chlorophyll solution, the path of the beam is marked by a brilliant red light. It was then noticed by Herschel[2] that when the sun's rays fall upon a dilute solution of sulphate of quinine the surface of the liquid, where the light falls, exhibits a bright blue colour, which penetrates to a small distance within the liquid. Herschel found that this blue light was confined to the surface stratum on which the light fell, and that light which had once produced this effect was incapable of developing it again in other solutions of the same substance.

Thus if the pencil of solar light, which passes through a cell containing sulphate of quinine, be transmitted through a second cell of the same solution, there will be no development of the blue shimmer at the surface of the second cell where the light enters it. That part of the solar beam which developed the blue colour in the first solution has been absorbed there, and the transmitted pencil contains no rays competent to excite it in the second. If the light be concentrated, by means of a lens, to a point in the interior of the liquid, the blue shimmer may mark the whole path of the beam, and be not merely restricted to a thin surface layer. The same appearances are exhibited

[1] Brewster, Eighth Report, *Brit. Assoc.* and *Phil. Mag.* 1848. [2] Herschel, *loc. cit.*

in the greenish surface colour of canary glass (coloured with oxide of uranium), and in the bluish surface tint so frequently observed in some kinds of ordinary paraffin oil. Fluorescence is also finely developed in solutions made from the bark of the horse chestnut.

Sir G. G. Stokes,[1] to whom the whole explanation of this phenomenon is due, has shown that it is exhibited in a greater or less degree by many substances, including white paper, bone, ivory, cotton wool, etc. In reflecting on the possible explanation of Herschel's observations it occurred to him that the blue light dispersed in the solution of quinine might not be the blue rays of the incident beam, but might be due to the degradation of the rays of higher refrangibility, which are mostly invisible, and the results of experiment completely verified the speculation.[2]

To find what rays are most effective in producing fluorescence in any substance, it is only necessary to place a piece of it in the various parts of a pure spectrum of solar (or electric) light. The spectrum should be formed with prisms and lenses of quartz, as this material transmits the ultra-violet radiations very freely. When the substance is moved from the red end of the spectrum towards the violet it is found that at a certain point it begins to glow with colour, always of a refrangibility lower than that which falls upon it. Thus when a slip of uranium glass is so treated there is scarcely a trace of colour exhibited until it is moved up to the blue part of the spectrum; here it begins to glow with its characteristic yellowish light, and the effect persists as it is moved through all the higher colours, and even to a considerable distance beyond the limits of the violet. The ultra-violet waves, which are too short to affect the sense of sight, are absorbed by the uranium glass, and afterwards emitted as rays of lower refrangibility, sufficiently slow to be detected by the eye.

Solutions of chlorophyll, sulphate of quinine, and other liquids may be examined in the same manner by enclosing them in a small test-tube. When a test-tube containing sulphate of quinine is held in the red end of the spectrum nothing strange is observed. In the red it looks red, and in the green it appears green, but in the blue and violet a marked change occurs in the appearance. The pale blue shimmer exhibited in the sun's rays begins to show itself, and increases as the solution is moved towards the end of the spectrum, remaining visible even beyond the limits of the violet. The red and green rays are inactive; it is the rays of higher refrangibility from the violet extremity of the spectrum that operate on the sulphate of quinine.

[1] Stokes, *Phil. Trans.* 1852-53.

[2] [Compare Ex. 8, Art. 48.]

If the blue fluorescent light emitted by the quinine be examined in a spectroscope it will be found to contain rays from various parts of the spectrum. It is not a homogeneous blue light. All its constituents, however, are of lower refrangibility than the violet or ultra-violet which excited them.

In chlorophyll the greater part of the fluorescent effect is produced by the visible light of the spectrum, and not so much by the ultra-violet, as in the case of sulphate of quinine. Held in the red end of the spectrum, the chlorophyll glows with a deep red. It has already been stated that the spectrum of solar light transmitted through a solution of chlorophyll exhibits a marked absorption band in the red between the lines B and C. The fluorescent light emitted by the solution when held between these lines is found, on analysis, to be of lower refrangibility than the light absorbed. As the test-tube is moved up the spectrum the red glow grows fainter, but again rises in intensity as each absorption band is crossed. In the blue and violet the red glow is more continuous, and at the end of the violet it assumes a brownish tint, due to the presence of some green in the fluorescent light.

**292. Application—Study of the Ultra-violet Spectrum.**—One important application of fluorescent substances arises in the means they afford us of mapping the solar spectrum beyond the limits of the violet. This remarkable prolongation of the solar spectrum beyond the violet was first noticed by Sir John Herschel in throwing the spectrum on turmeric paper. Herschel attributed this prolongation, which was yellow, to a peculiar reflecting power of the paper.

Casting the complete solar spectrum on the surface of a transparent fluorescent substance, various effects of the different colours may be observed simultaneously. At the red end there is in general no fluorescent effect. Towards the blue and violet, and for considerable distances beyond the confines of the violet, the fluorescent shimmer is emitted from the surface layer, and dark lines are visible in some parts, indicating absorption bands in the solar spectrum beyond the violet similar to the Fraunhofer lines which are presented in the coloured part. The waves at these parts do not exist in the solar beam when it reaches us. They have been absorbed in the solar or terrestrial atmosphere, or else they have not been emitted by the sun at all.

**293. Phosphorescence.**—The fluorescent light which we have been discussing is emitted only while the substance producing it is illuminated. On intercepting the light which falls upon a solution of sulphate of quinine the blue shimmer from its surface layer disappears at once. Some substances, however, especially the sulphides of barium, strontium, and calcium, persist in emitting light even after the incident beam has been cut off—that is, they shine in the dark for some time after exposure to the light. This phenomenon is called *phosphorescence*. It has, however, nothing to do with the luminosity of phosphorus, which is due to slow oxidation. It is merely fluorescence lasting after the exciting cause has been removed. The light absorbed during exposure is emitted as light of lower refrangibility, and continues to be emitted by many substances for long or short intervals after they are removed from the light and brought into a darkened chamber. The duration of the phosphorescence after the incident light has been cut off is, in most substances, so short that it is impossible to detect it without some specially contrived method of investigation. To pursue this inquiry, Becquerel invented an ingenious *phosphoroscope*, by means of which we can determine with considerable accuracy the duration of the phenomenon after the direct light has been cut off. With this apparatus the existence of phosphorescence to a greater or less degree has been detected in most substances.

## Theories of Dispersion

**294. Dependence of the Velocity on the Wave Length.**—We owe to Newton the important discovery that the light of the sun is composite, that it consists of a system of simple colours, or, as we say now, of a great variety of waves of different periods. The separation of solar light into its constituent simple colours was effected by Newton with the aid of a prism of glass. The different colours are deviated by different amounts in passing through the prism, and the transmitted light, when received on a white screen, paints on it a coloured band or spectrum. The property which transparent substances possess of thus separating the various constituents of white light is called *dispersion*, the inequality in the refraction of the various colours leading to their separation or dispersion from one another.

Now the teaching of the wave theory, and of the emission theory also, is that rays which are refracted by different amounts on entering any medium must travel through that medium with different velocities, the relation between the absolute refractive index and the velocity for any ray being, according to the wave theory,

$$\mu = \frac{v_0}{v},$$

where $v_0$ is the velocity in free space—that is, in the ether. This velocity is, as far as we know, the same for light of all colours. A condition which may therefore be introduced into any theory of dispersion is that it does not exist in free space, and is a phenomenon arising from the interaction of the ether and matter.

The constancy, or approximate constancy, of $v_0$ for all waves ranging from the red to the violet at least, is inferred from considerations such as the following. Certain stars exhibit rapid changes in brightness and are called variable stars. One of these, Algol, passes in three and a half hours from one of the second to one of the fourth magnitude in brightness. Whatever be the cause of these changes, whether it be due to eclipse or otherwise, they ought to be accompanied by corresponding exhibitions of colour, if the various colours travel through space with different velocities. Thus if the red light travels faster than the violet in interstellar space, as it certainly does in glass and common transparent substances, then when the star is growing faint it should be coloured blue or violet, and when it is growing bright it should appear red. No trace of a coloured tint has ever been observed, although light requires several years to reach us from this star. If the difference in the velocities of the red and violet amounted to the one hundred thousandth part of the value of either, that is a variation of ·001 per cent, an interval of at least an hour would elapse between their times of arrival and the corresponding changes in tint should be observed. We may therefore conclude that the waves corresponding to the various colours of the spectrum traverse the free ether with the same velocity, but it is not to be assumed that waves of every length pass with the same velocity. The foregoing observations indicate only an equality, or an approximate equality, between the velocities of the very limited sets of waves which affect the eye. The long waves far below the limits of the red may travel with velocities very different from that with which the short ultra-violet waves are propagated.

In the case of sound notes of all pitches travel with equal velocities, and the common velocity deduced by theory is a result of the supposition that the length of the wave is vastly greater than the displacements, or sphere of action, of the vibrating molecules. When the wave length is very small, as in the case of light, the conditions are very different, and we cannot assume that the velocity of propagation will be independent of the time of vibration. Thus in the case of a stretched string, vibrating transversely, the velocity of propagation will depend upon the wave length if the stiffness of the string be taken into account.

**295. Cauchy's Formula.**—Starting from these principles, Cauchy [1] showed that the velocity will in general be a function of the wave

[1] Cauchy, *Nouveaux Exercices de Mathématique*, 1835.

length. By proceeding to a higher order of approximation he arrived at the relation

$$v^2 = a + \frac{b}{\lambda^2} + \frac{c}{\lambda^4} + \frac{d}{\lambda^6} + \ . \ . \ .$$

which expresses the refractive index in the form

$$\mu = A + \frac{B}{\lambda^2} + \frac{C}{\lambda^4} + \ . \ . \ .$$

where A, B, C, etc., are constants depending on the nature of the medium and diminishing rapidly in magnitude as we proceed to the higher terms.

The formula shows that waves of short period are more highly refracted than the waves of longer period, and that the latter approach a limited index $\mu = A$. Within the limits of the visible spectrum it represents, when taken to three or four terms, the results of experiment very accurately, but fails to represent the facts when applied to the dark radiations beyond the red, as shown by the experiments of Prof. Langley.[1] In media, however, which exhibit anomalous dispersion it is not true that shorter waves possess the higher refractive indices, and the formula takes no account of this class of phenomena, which point to an intimate connection between the production of dispersion and the existence of absorption, the former being perhaps in all cases the result of the latter.

In Cauchy's method of investigation the assumption is that the distances between the molecules are comparable with the wave length, but in the modern methods which consider the direct interaction of the ether and matter, it is the periodic time of the wave which is supposed to agree with, or approach, the period of free vibration of the molecule, so that absorption takes place.

**296. Briot's Formula.**—The work of Cauchy was taken up and treated in a more general manner by Briot.[2] Taking into account more directly the interaction of the ether and the matter molecules, he obtained the formula,

$$\frac{1}{\mu^2} = \kappa\lambda^2 + A + \frac{B}{\lambda^2} + \frac{C}{\lambda^4} + \ . \ . \ .$$

which consists of two parts, one a series similar to that of Cauchy, and the other the term $\kappa\lambda^2$, which depends on the direct action between the ether and the matter.

The matter is supposed to affect the ether in two ways. Firstly, by modifying its distribution in the body, and this gives rise to the Cauchy series; secondly, by exercising a direct action on its motion

[1] S. P. Langley, *Ann. de Chimie et de Physique*, tom. ix. p. 496, 1886.

[2] Briot, *Essai sur la Théorie mathématique de la Lumière*, Paris, 1864.

when vibrating, and this introduces the term $\kappa\lambda^2$. This formula represents to a high degree of accuracy the dispersion of the luminous rays, and also agrees fairly with Langley's investigations in the infra-red.[1]

**297. Anomalous Dispersion.**—To account for the existence of anomalous dispersion, theories, based on the mutual reaction between the ether and matter, have been developed by Helmholtz, Ketteler, Lommel, etc. The forces taken into account are the elastic reactions of the ether, the elastic reactions in the matter, and the mutual forces arising from the interaction of the ether and the matter. Each element of the ether is subject to forces arising from its own elastic reactions and to forces due to the action of the matter, while each matter molecule is subject to the elastic reactions of the matter, the action of the ether, and to frictional forces arising from the adjacent matter. Expressing these conditions, differential equations are formed for the motion of the ether and also for the motion of the matter. When the coefficient of absorption is small, so that its square may be neglected, Wüllner has shown that Helmholtz's formula for the motion of the ether may be reduced to the form

$$\mu^2 - 1 = P\lambda^2 + \frac{Q\lambda^4}{\lambda^2 - \lambda_1^2}.$$

This formula corresponds to a single absorption band produced by a single system of molecules. To account for the existence of several bands, as many systems of molecules, vibrating in corresponding periods, must be introduced, and as many systems of equations integrated.[2]

[1] Ketteler finds that the formula

$$\mu^2 = -\kappa\lambda^2 + a^2 + \frac{D\lambda_m^2}{\lambda^2 - \lambda_m^2}$$

represents the facts very accurately. From Langley's experiments on *sel gemme* he determines

$$\kappa = 0{\cdot}000858, \quad a^2 = 2{\cdot}32883, \quad D = 1{\cdot}1410, \quad \lambda_m^2 = 0{\cdot}01621.$$

(*Journal de Physique*, 2e, tom. vii., 1888.)

[2] [See Lord Kelvin, *Phil. Mag.*, March 1899, on the application of Sellmeier's dynamical theory to the D lines of sodium vapour.]

# CHAPTER XIX

## EXPERIMENTAL DETERMINATIONS OF THE VELOCITY OF LIGHT

**298. Introduction.**—The first attempt to determine the velocity of propagation of light was made by Galileo. The principle of the method employed was as follows. Two observers, A and B, are situated at a distance, and each is furnished with a lamp which can be quickly screened or uncovered. If A uncovers his lamp it will be seen after a certain interval by B, and this interval will measure the time occupied by the light in travelling from A to B. If B now uncovers his lamp it will be seen after an equal interval by A. Hence if B uncovers his lamp at the instant he perceives that of A, then A will perceive that of B at a time after he uncovered his own, equal to twice the time required by light to traverse the distance between A and B, together with a small interval of time depending on B, namely, the time required by B to perceive the flash of A's lantern and uncover his own. This *personal* interval becomes of such weight (when compared with the very small time required by light to traverse such distances as those which could be employed in an experiment of this character) that if any definite result had been obtained it would have been altogether worthless. The fundamental principle of this method, however, is correct, and is the same as that on which one of the most celebrated of modern methods is based, namely, Fizeau's, or that which may in general be termed the *eclipse method.*

It is interesting to notice how the academicians might have obtained an estimate of the velocity of light by attending to the conditions on which the delicacy of their method depended. In the first place, having determined that the velocity of light, if not infinite, was certainly very great, it becomes a matter of prime importance to eliminate the personal interval arising from the slowness of perception and movement of the second observer B. At the instant the light from the first station reaches the second, the conditions of the problem require that a return beam of light shall leave the second station and travel to the first. The most obvious way of securing this is to replace the second observer B by a reflecting surface, such as a polished mirror, so that when A uncovers his lantern the light issuing from it may fall upon the mirror and be reflected back at the second station without loss of time. Thus when A uncovers his own lantern he may observe its image by reflection in the mirror at the second station, and the interval between the uncovering of the light and the perception of the image is the quantity to be estimated.

The whole problem is now reduced to devising an accurate method of estimating this interval. One obvious source of error is a personal one attending the observer, and arises in the uncertain estimate of the interval between the muscular effort in removing the screen and the instant at which he considers he has perceived the return flash of light. To eliminate this the screen used to uncover the light should in its motion be employed to cut off the return rays. For example, let the lantern be covered, and let the eye of the observer be placed close to it and directed so as to look into the distant mirror. Then if the screen be suddenly moved so as to uncover the lantern and cover the eye of the observer, the return flash will not be received by the eye if the time occupied in moving the screen from the lantern to the eye is less than the time required by light to travel to the distant mirror and return.

The problem is now reduced to the devising of some mechanical method of moving a screen, or a train of screens, across the field of view so that the source of light and the eye of the observer shall be covered and uncovered in succession. This was first done in 1849 by Fizeau by using a toothed wheel, which rotated in the field of view in such a way that the opaque teeth passed before the eye and intercepted successively the light emerging from the source and that returning to the eye after reflection in the distant mirror. This method is explained in the following article.

### Fizeau's Method

**299. General Principles of the Method.**—In the method adopted by M. Fizeau[1] a beam of light from a source S (Fig. 216) is introduced through a collimator fixed to the side of a telescope, and after reflection at the surface of a plate of parallel glass inclined at 45° to the axis of the tube, the pencil comes to a focus at a point F, which is the principal focus of the object-glass of the telescope. It

[1] H. Fizeau, *Comptes Rendus*, tom. xxix. p. 90, 1849.

follows therefore that the light diverging from F will emerge from the telescope in a parallel beam. This beam, after traversing a distance of three or four miles, falls upon a lens L, and is thereby focussed on a reflector R, which is part of the surface of a sphere having its centre of curvature at the centre of the lens L.

On account of this arrangement it follows that the beam of light, as a whole, after reflection at the surface of the mirror, emerges from the lens L in a parallel pencil, and, retracing its former path,[1] is again brought to a focus at the point F. It then diverges from F and falls upon the inclined glass plate, where it is in part reflected and in part transmitted. The transmitted portion is received by an eyepiece E, and enters the eye of the observer, so that an image of the source S is seen in the field of view.

The eyepiece E combined with the object-glass at the observing station constitutes a telescope, which may be referred to as the "observing" telescope, and similarly the apparatus RL at the distant station

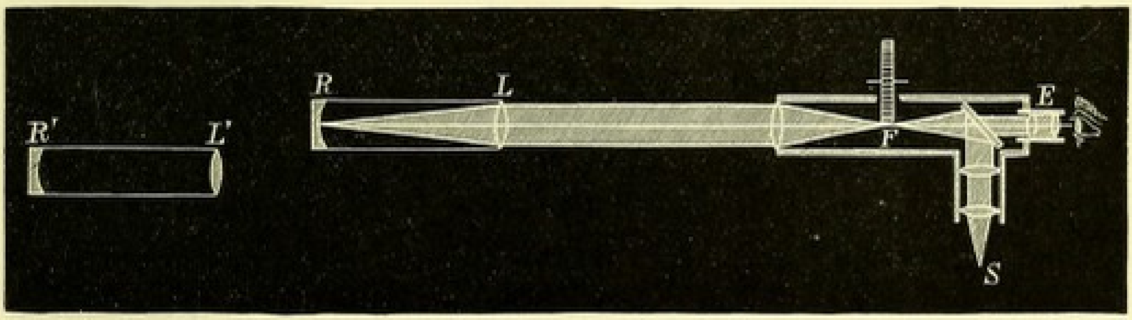



Fig. 216.

may be referred to as the reflecting collimator. When the apparatus is properly adjusted these two pieces should be directed towards each other in such a way that an image of the object-glass of each is formed in the principal focus of the other. This might be effected before the mirror R is placed in position by attaching an eyepiece to the tube RL so as to form a true telescope, whose axis can be directed as required towards the observing station. This being done the mirror R has to be placed in position. In order to do this with precision the following device was resorted to by Messrs. Young and Forbes in the experiments described subsequently (Art. 302). The mirror was attached to a cap which screwed on to the end of the tube RL, and an exactly similar cap was fitted with a piece of ground glass, so that the glass in one cap occupied the same position as the mirror in the other. The ground glass cap was first screwed on to the tube RL, so

[1] It is to be observed that a ray entering one half of the lens L is returned after reflection at R through the other half, so that the path of the reflected *ray* does not coincide with that of the incident.

that the image was received upon the glass in proper focus, the number of turns of the screw required to effect this being noted. The glass cap was then removed, and the mirror cap screwed on through the same number of turns. The centre of the mirror must now be brought into coincidence with the centre of the lens L, and this was effected by means of three screws at the back of the mirror. To test this adjustment a tube about one foot long was placed on the collimator so as to project in front of the object-glass, and a small ring was placed at the end of this tube, the ring being supported at the centre of the tube by three strips of metal. On looking through this ring the observer ought to see an image of his eye when the adjustment is perfect, and the screws at the back of the mirror were altered until this was the case.

To direct the reflecting collimator RL so that its axis shall point to the observing telescope, it was found most convenient to look through its object-glass L at the mirror R (the head of the observer being kept out of the path of the light as far as possible). An image of the distant station was then seen in the mirror, and the direction of the axis was changed until the light coming from the observing telescope was seen in the centre of the mirror.

A toothed wheel, connected to a clockwork driven by weights, so that it could be set in uniform and rapid rotation round an axis parallel to the axis of the telescope, is placed with its edge at F, so that as it rotates the light is alternately intercepted and allowed to pass between its teeth. Let us suppose the wheel to be at rest, and that F is situated in the space between two teeth. In this case the beam is allowed to pass, and a bright image is seen in the field of view. But if F falls upon a tooth, the light reaching it from S will be reflected directly back into the eye. To avoid this the teeth may be blackened, or bevelled so as to reflect the light against the sides of the telescope which are also blackened. If this is secured there will be no illumination in the field except when the light passes through a tooth space and returns after reflection at the distant mirror.

Now if the wheel rotates very slowly the image in the field of view will appear and disappear successively as the spaces and teeth pass before F, but if the speed be increased so that several teeth pass per second, the succession of brightness and darkness will be so rapid that, owing to the persistence of the visual impression, a permanent image will be seen. Hence if the angular width of a space be $\alpha$ and the width of a tooth be $\beta$, so that the combined width of a tooth and space is $\alpha+\beta$, it will follow that if the intensity of the image seen when the wheel is at rest be $I_0$, its intensity when the wheel is rotating slowly (but rapidly enough to cause a continuous impression) will be

$$I = \frac{\alpha}{\alpha + \beta} I_0.$$

If the velocity of light is infinite the illumination of the image will remain constant for all speeds, but if it is propagated in time then I will depend upon the speed of rotation. Some of the light transmitted through a space will in returning fall upon the adjacent tooth and be intercepted, and if the speed be great enough, so that when the light returns a tooth has moved into the position previously occupied by the space, then all the returning light will be intercepted, provided the teeth be at least as wide as the spaces, and complete extinction will be effected.

What occurs, therefore, is that at first a bright image is observed, which diminishes in brightness as the speed of rotation is increased, and is finally extinguished if the width of a tooth be equal to or greater than that of a space. If $\beta$ is greater than $\alpha$ the light will remain eclipsed until the speed is sufficiently increased to remove the obstructing teeth and bring the spaces into position for the returning light. The image now reappears and gradually grows in brightness as the speed is raised. It reaches a maximum and then fades away again into darkness, and so on in succession for higher and higher speeds. If $\alpha = \beta$, or the teeth and spaces are equal in width, then for certain particular speeds the light will be completely eclipsed, but will reappear for speeds either less or greater. If $\beta$ is less than $\alpha$ the light will never be wholly eclipsed, but will merely fall to a minimum and rise again to a maximum in alternate succession.

In Fizeau's experiments the teeth and spaces were of equal width, each being a quarter of a degree, so that the wheel possessed 720 teeth. In this case, if D be the distance between the toothed wheel and the reflector, $v$ the velocity of light, and T the time occupied in traversing the distance 2D, we have

$$T = \frac{2D}{v}.$$

In this time the wheel will have turned through an angle $\omega T = 2\pi n T$, if $n$ is the number of revolutions per second. Consequently, if $\alpha$ be the angular width of a space, the first eclipse will occur when

$$\alpha = 2\pi N_1 T = 4\pi N_1 D/v,$$

where $N_1$ is the speed of the revolution when the first eclipse occurs. If the speed increases the image will reappear and grow in intensity till it reaches maximum brightness at the speed

$$2\alpha = 4\pi N_2 D/v.$$

As the speed increases further the illumination will wane, and a second eclipse will occur when

$$3\alpha = 4\pi N_3 D/v.$$

Similarly the $p$th eclipse will occur at the speed $N_{2p-1}$, where

$$(2p-1)\alpha = 4\pi N_{2p-1} D/v,$$

from which we have

$$v = \frac{4\pi D N_{2p-1}}{(2p-1)\alpha} = \frac{4m D N_{2p-1}}{2p-1},$$

where $m = \pi/\alpha$ is the number of teeth contained in the wheel. Reckoning the first maximum brightness as that which occurs after the first eclipse, it is clear that the $p$th maximum will occur at the speed $N_{2p}$ given by the equation

$$2p\alpha = 4\pi N_{2p} D/v,$$

from which we have a corresponding formula for $v$.

Fizeau endeavoured to determine the speeds at which the successive eclipses occurred, but this was a matter of extreme difficulty and uncertainty. In the first place, the intensity of the light returning to the telescope is greatly weakened by transmission through the apparatus and by reflection at the glass plate G, so that the image seen is necessarily faint even when at its maximum brightness. It is again rendered less distinct by the extraneous illumination in the field of the telescope, caused by reflection from the teeth of the wheel. For when the wheel rotates, the light when not passing between the teeth is reflected from them back into the field of view, and causes a general illumination of the whole field. It is therefore very desirable to suppress this reflected light as much as possible, and to this end Messrs. Young and Forbes in repeating the experiment bevelled the teeth so that the light reflected from them fell upon the blackened sides of the telescope. They also smoked the wheel itself, so as to diminish its reflecting power as much as possible.

But even when the distinctness of the image is secured it is very difficult to decide at what instant it is completely extinguished, it being much more easy to say when two images seen simultaneously are equally intense than when any image of varying intensity has reached its maximum or minimum brightness. Determinations deduced from direct observations of the speed at maximum or minimum brightness are consequently attended with considerable uncertainty.

Working with a distance $D = 8633$ metres, Fizeau found that the first eclipse occurred when the speed of revolution was 12·6 turns per second, and the final result for the velocity was put down at 70,948 leagues of 25 to the degree. This is taken to represent a velocity of about 315,000 kilometres per second.

**300. Cornu's Experiments.**—In 1874 M. Cornu[1] made a determination of the velocity of light by Fizeau's method with greatly improved apparatus. Accuracy is very difficult to attain in the method of observation adopted by Fizeau, since it is almost impossible to determine when the image is exactly eclipsed. To evade this difficulty M. Cornu placed the mechanism of the toothed wheel in electrical connection with a chronograph, so as to mark every hundred revolutions. At the same time a clock marked seconds, and tenths of seconds were recorded by means of a vibrating spring. The observer had also under his control a key by means of which he could record any instant at which he wished to know the velocity. From the chronographic record the speed and rate of change at every instant could be obtained. Allowing the speed to increase, the illumination will sink to a certain value and the speed is then signalled. After complete extinction the image will reappear, and when it attains the former brightness the speed is again signalled. The speed corresponding to zero brightness is the mean of these two, and is found from the chronographic record.

In M. Cornu's experiments the distance D between the two stations was nearly 23 kilometres, so that he was enabled to observe eclipses up to the thirtieth order—that is, to make 15 teeth of the wheel pass before the flash from the distant mirror returned.

The method, however, is not very desirable, in that it is not capable of delicate measurement in regard to the quantity directly estimated—namely, the brightness of the image. For the eclipses are not sudden phenomena occurring at well-marked speeds, but are so gradual that it is difficult to say precisely when they occur, and even in the method adopted by M. Cornu to evade this difficulty there must still remain considerable uncertainty. His investigations, however, surpass in reliability and comprehensiveness anything which was previously attempted.

The final value obtained for the velocity of light was 300,330 kilometres per second in air, and this corresponds to 300,400 kilometres in vacuo. This result is, however, somewhat too high, for recent experiments have established that the velocity is undoubtedly less than 300,000,000 metres per second.[2]

[1] Cornu, *Annales de l'Observatoire de Paris* (*Mémoires*, tom. xiii. 1876).

[2] Assuming that the velocity determined by the toothed-wheel method ought to be the same as that determined by the revolving mirror. [At the present date a new determination of the velocity of light is being made at Nice Observatory by M. Perrotin using Fizeau's original apparatus as modified and used by Cornu (*Comptes Rendus*, tom. cxxxi. pp. 731-734). The preliminary results of 1500 measures indicate a velocity of $299{,}900 \pm 80$ km., which agrees far better with the value derived from the revolving mirror method than with Cornu's determination.]

**301. Brightness of the Image at any Speed.**—If the angular width of the spaces and teeth be $\alpha$ and $\beta$ respectively, and if the intensity of the image when the wheel is at rest be $I_0$, its intensity when the wheel is rotating slowly will bear to $I_0$ the ratio $\alpha : (\alpha + \beta)$, for this is the ratio of the quantity of light allowed to pass per second when the wheel is rotating slowly to the quantity when it is at rest or altogether removed. But if the wheel rotates through an angle $\epsilon$ while the light travels over the double journey between the two stations, the space available for the transmission of the returning light will be reduced from $\alpha$ to $\alpha - \epsilon$, and the intensity of the image will be

$$I = \frac{\alpha - \epsilon}{\alpha + \beta} I_0,$$

where $\epsilon = \omega T = 2\pi n T$, when the wheel is making $n$ revolutions per second.

If the number of teeth be $m$ we have $m(\alpha + \beta) = 2\pi$, and therefore $\epsilon = mnT(\alpha + \beta)$, so that if $\alpha/(\alpha + \beta) = k$ we obtain

$$I = (k - mnT)I_0.$$

Hence the intensities are represented by the ordinates, and the speeds by the corresponding abscissæ of the right line

$$y = (k - mTx)I_0,$$

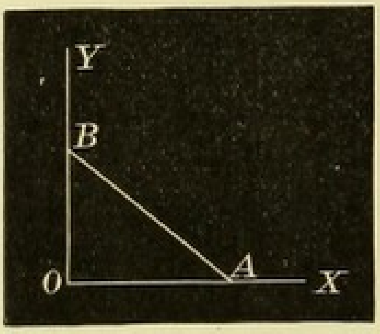


Fig. 217.

which makes an intercept $OB = kI_0$ on the axis OY (Fig. 217) corresponding to zero (very slow) speed[1] and maximum intensity, and an intercept $OA = k/mT$, which corresponds to the first extinction. Three cases are presented according as $\alpha$ is equal to, less than, or greater than $\beta$.

If $\alpha = \beta$, that is, if the teeth and spaces are equal in width, the initial intensity (for a low speed) is $kI_0 = \frac{1}{2}I_0$. As the speed increases the intensity falls in proportion, and complete extinction occurs at the speed

$$N_1 = \frac{k}{mT} = \frac{1}{2mT} = \frac{v}{4mD}.$$

As the speed increases beyond this value the intensity rises proportionately, and again attains the maximum value $kI_0 = \frac{1}{2}I_0$ at the speed $N_2 = 2N_1$. A further increase of speed diminishes the intensity, and a second extinction occurs at the speed $N_3 = 3N_1$, and so on in succession. The variation of the intensity with the speed is represented by the broken line $BA_1B_1A_2$, etc. (Fig. 218).

[1] That is, only fast enough to produce persistence of the visual impression.

If $\alpha < \beta$ or $k < \frac{1}{2}$, then the first extinction occurs at the speed $\epsilon = \alpha$, or $OM = k/mT = kv/2mD$ (Fig. 219). The intensity will now remain zero till $\epsilon = \beta$, which occurs at the speed $ON = (1-k)v/2mD$. The brightness again rises to a maximum $kI_0$ at $B_1$ corresponding to a speed $\epsilon = \alpha + \beta$ or $N = v/2mD$. The arithmetic mean of the speeds OM and ON, corresponding to the first extinction and first reappearance, is equal to the speed at which the first eclipse would occur if the teeth and spaces were equal, and each $\frac{1}{2}(\alpha+\beta)$ in angular width. This speed, corresponding to $A_1$ (Fig. 219), the first

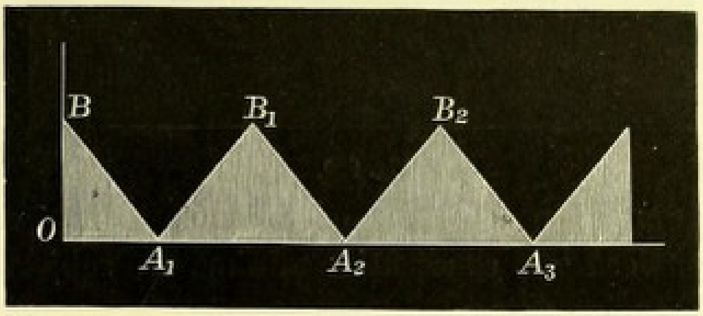



Fig. 218.

*central* extinction, is $\frac{1}{2}(OM + ON) = v/4mD = N_1$. The speed corresponding to $B_1$, the first maximum brightness, is $2N_1$; that at the second central extinction, $A_2$, is $3N_1$, and in general the maximum brightnesses occur at the even multiples of $N_1$, and the central extinctions at the odd multiples.

Since $mT = 2mD/v = 1/2N_1$, the expression for the intensity at any speed $x$ before the first eclipse is

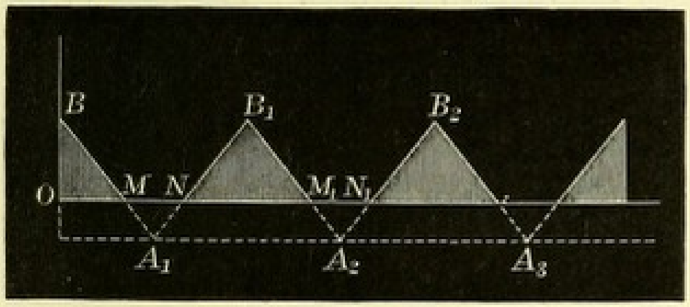



Fig. 219.

$$I = \left(k - \frac{x}{2N_1}\right)I_0,$$

and since the intensity must be the same when the speed is increased from $x$ to $n = 2pN_1 \pm x$, it follows that between the $p$th and $(p+1)$th eclipses we have $(\pm x = 2pN_1 - n)$

$$I = \left(k \pm \frac{2pN_1 - n}{2N_1}\right)I_0 = \left\{k \pm \left(p - \frac{n}{2N_1}\right)\right\}I_0$$

the positive sign applying to the case in which the brightness is decreasing and the negative sign to that in which it is increasing. Thus between the $p$th and $(p+1)$th eclipses, when the brightness is increasing $(n < 2pN_1)$,

$$I = I_0\left(k - p + \frac{n}{2N_1}\right) \qquad \text{(brightness increasing)},$$

and when the brightness is decreasing $(n > 2pN_1)$

$$I = I_0\left(k + p - \frac{n}{2N_1}\right) \qquad \text{(brightness decreasing)}.$$

When $\alpha > \beta$, or $k > \frac{1}{2}$, the image is never totally eclipsed but sinks from a maximum $kI_0$ to a minimum determined by the speed $\epsilon = \beta$, and the corresponding minimum intensity is given by the equation

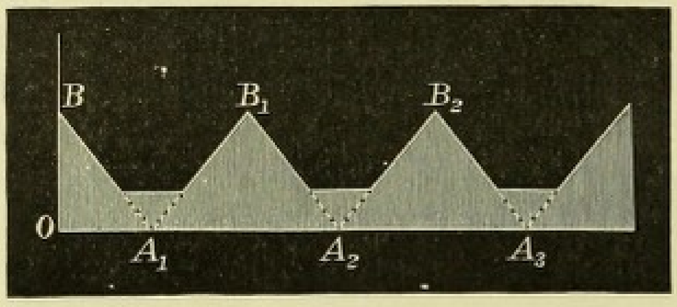



Fig. 220.

$$I = \frac{\alpha - \beta}{\alpha + \beta} I_0 = (2k - 1)I_0.$$

The intensity remains stationary, as represented in Fig. 220, till the speed attains the value $\epsilon = \alpha$, from which it will increase and attain a maximum when $\epsilon = \alpha + \beta$, and so on. The variations of the intensity between two consecutive minima are subject to the same laws as in the foregoing cases.

**302. Experiments of Young and Forbes.**—An elaborate series of experiments on the velocity of light was executed by Dr. J. Young and Professor G. Forbes,[1] using a modified form of Fizeau's apparatus. The chief novelty of their method is the introduction of a second reflecting telescope R′L′ (Fig. 216) situated behind the first and nearly in the same line, so that two images were seen close together (not $\frac{1}{100}$ of an inch apart) at the edge of the revolving wheel. The ratio of the distances of the two reflectors from the toothed wheel was 12 : 13. Denoting these distances by D and D′, and the speeds at which the corresponding images are first eclipsed by $N_1$ and $N'_1$, it is clear that if D′ be greater than D, then $N_1$ will be greater than $N'_1$, so that the image from the farther reflector will be first eclipsed and also the first to reappear. What happens then is that as the speed of the wheel is increased from rest, one image I′ fades more rapidly than the other I, till the former is extinguished and the latter remains alone in the field, if the width of the spaces is less than that of the teeth. One of two things may now occur. The image I′ may reappear before the other I is completely extinguished, or the latter may be eclipsed before the former reappears, so that both images are extinguished. In the former case I′ on reappearance will increase in brightness with the speed, and I approaching eclipse will continually diminish, so that a speed will be reached at which the two are equal in brightness; we have then $I = I'$, or

$$I_0\left(k - \frac{n}{2N_1}\right) = I'_0\left(k - 1 + \frac{n}{2N'_1}\right).$$

The speed being increased, I′ will increase and I diminish to zero, and after eclipse will reappear while I passes through its maximum, and

[1] Young and Forbes, *Phil. Trans.* pt. i. 1882.

begin to diminish. As I′ diminishes I will increase, and equality will again be reached, when we have

$$I_0\left(k-1+\frac{n}{2N_1}\right)=I'_0\left(k+1-\frac{n}{2N'_1}\right),$$

and in general, if equality occurs between the $p$th and $(p+1)$th eclipses of both, we have

$$I_0\left(k-p+\frac{n}{2N_1}\right)=I'_0\left(k+p-\frac{n}{2N'_1}\right).$$

If equality be again established after the next eclipse of I′—that is, after the $p$th eclipse of I and the $(p+1)$th eclipse of I′—since I is now decreasing and I′ increasing, if the speed be $n'$, we have

$$I_0\left(k+p-\frac{n'}{2N_1}\right)=I'_0\left\{k-(p+1)+\frac{n'}{2N'_1}\right\}.$$

Let $I'_0=\rho I_0$ and $N'_1=gN_1$, then we have $g=D/D'$, and the foregoing equations become

$$k-p+\frac{n}{2N_1}=\rho\left(k+p-\frac{n}{2gN_1}\right),$$

$$k+p-\frac{n'}{2N_1}=\rho\left\{k-(p+1)+\frac{n'}{2gN_1}\right\},$$

which by subtraction give

$$2p+\rho(2p+1)=\frac{1}{2N_1}(n+n')\left(1+\frac{\rho}{g}\right).$$

Now if the distances D and D′ be so chosen that their ratio $g$ is equal to $2p/(2p+1)$, we have

$$2p\frac{g+\rho}{g}=\frac{(n+n')}{2N_1}\frac{g+\rho}{g},$$

or

$$N_1=\frac{n+n'}{4p},$$

which gives $N_1$ in terms of observed quantities, and hence we have the velocity of light in the form

$$v=4mDN_1=\frac{mD(n+n')}{p}.$$

In the experiments of Young and Forbes the observations were made at the 12th equality, and $g$ was almost exactly equal to 12/13. The value of $v$ deduced was 301,382,000 metres per second. The source of light employed was a Siemens electric lamp, and the speed was registered electrically. The distances D and D′ were 3·18845 and 3·44928 miles respectively.

Any want of achromatism in the lenses will impart a corresponding colour to the image, and atmospheric absorption will produce a similar effect. Messrs. Young and Forbes observed that one of the images was in general red and the other blue, but closer examination proved that when the brightness of either was increasing its colour was red, and when its brightness was fading the colour was blue. This they concluded to indicate a difference in speed of the rays of higher and lower refrangibility. Thus if the blue rays travel faster than the red, then the red rays will be eclipsed at a lower speed than the blue, so that as the speed of the wheel is increased, the rays from the red end of the spectrum will be first eclipsed, and the fading image will appear blue. On the other hand, when the image is reappearing after eclipse, the rays which travel slowest will be the first to gain admission through the adjacent tooth space, and the growing image will appear red.

Coloured image.

To test this point experiments were made with the red light and blue light from the spectrum (of the electric arc) formed by a prism, and from the average of the results the experimenters concluded that the blue rays travel about 1·8 per cent faster than the red. This difference is so great that, in the absence of other support, the effects observed have not been generally accepted as due to a difference in the velocities of the various rays, but it is surmised that the colouring is rather due to some extraneous cause not yet fully determined.

It is clear, however, that such a great difference as 1·8 per cent in the velocities of the red and blue rays should be detected in the other methods of estimating the velocity of light. For example, in Foucault's method, to be presently described, the image of the slit should be drawn out into an elongated spectrum, but no such colouring or elongation has ever been observed.

## Foucault's Method

**303. Foucault's Experiments.**—As early as 1834 Wheatstone[1] had employed a rotating mirror to determine the velocity of electricity and the duration of the electric spark. He further conceived that the same method might be used to determine the velocity of light, and test between the rival theories as to whether the speed of light was greater in the more refracting or less refracting media. The suggestion was taken up by Arago,[2] but it was not until 1850 that the experi-

[1] Wheatstone, *Phil. Trans.* p. 583, 1834.

[2] Arago, *Annuaire du Bureau des Longitudes pour* 1842, p. 287.

ment was designed in a form capable of giving accurate results by the ingenuity of M. L. Foucault.[1] The principle of the method is as follows :—

Solar light, transmitted through a rectangular aperture S (Fig. 221), falls upon an achromatic lens L, and afterwards upon a plane mirror R, which can be made to rotate rapidly round a vertical axis, the plane of the paper being supposed horizontal. A concave mirror M is fixed at a distance. The surface of this fixed mirror is spherical and its radius is equal to the distance RM, while its spherical centre is at R on the axis of rotation of the moving mirror. Let us first suppose the mirror R at rest, and so placed that the light reflected from it comes to a focus upon the fixed mirror M, and forms there a real image of the slit S. The pencil reflected from M returns along its former path, is reflected from R, traverses the lens a second time,

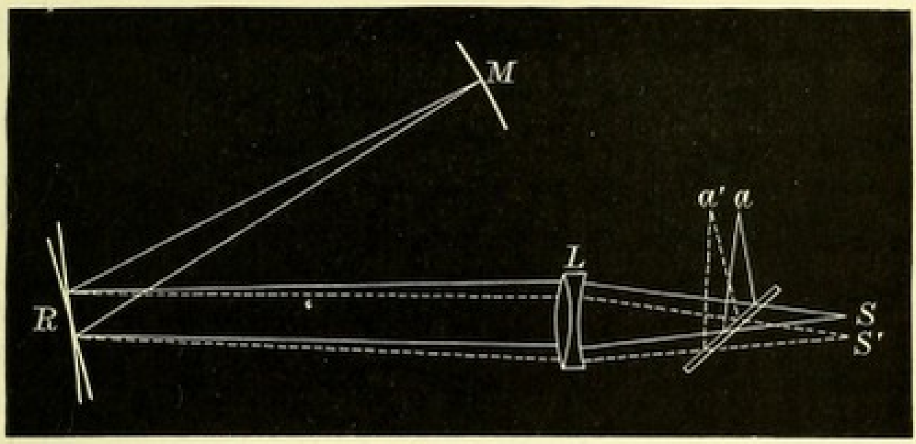



Fig. 221.

and comes to a focus at S, forming an image superposed on the slit. For the convenience of observation a plate of parallel glass is placed near S in the path of the beam of light, and inclined to it at an angle of 45°. The pencil reflected from M when returning to S meets the plate where it is in part reflected, and forms an image of S at $a$, which is observed through an eyepiece. A fine wire may be placed across the centre of the slit parallel to its length, that is vertically, so that the image at $a$ is crossed by a dark vertical line, over which the fibre of the eyepiece can be accurately [placed in making the measurements.

Let us now suppose that the mirror R is caused to rotate. As long as the light from R falls on M there will appear an illuminated image at $a$; and it is important to remark that on account of the curvature of M and its arrangement, as already described, the image $a$ remains fixed as R revolves, but when R turns round so that the reflected beam

[1] Foucault, *Comptes Rendus*, tom. xxx. p. 551, 1850 ; tom. lv. pp. 501, 792, 1862.

does not fall upon M there will be no illumination at $a$. If the rotation be very slow, brightness and darkness will alternately succeed each other at $a$, and when the rotation is sufficiently rapid the persistence of the visual impression leads to a permanent image appearing at $a$.

Brightness. The brightness of this image will obviously be much less than when the mirror is at rest, the ratio of the two being that of the arc M to a whole circumference. Hence by increasing the magnitude of M the brightness will be increased in the same ratio, and if in addition the revolving mirror be polished on both sides the illumination will be doubled.

Let T be the time required by the light to traverse and return along the distance $RM = D$, then $vT = 2D$. But during this interval the mirror R has turned through an angle $\omega T$, if its angular velocity be $\omega = 2n\pi$, where $n$ is the number of turns per second. The axis of the pencil returning through the lens to $a$ will consequently be rotated through an angle $2\omega T$, viz. twice the rotation of the mirror. The image $a$ will consequently be displaced to $a'$, and the image of S to S′, where $SS' = aa' = x$ suppose. The distance $x$ is measured by means of the micrometer attached to the eyepiece.

Now the light returning from M is reflected from R and appears to come from a point situated at an equal distance behind R, so that the pencils forming the images at S and S′ appear to come from sources $S_1$ and $S'_1$ behind R, wherefore $RS_1 = RS'_1 = D$, and the lines joining S and S′ to the optic centre of the lens pass through $S_1$ and $S'_1$ respectively. Denote the distances of the lens from the slit and revolving mirror by $a$ and $b$ respectively. Then since the angle SLS′ is very small, we have its circular measure

$$\frac{SS'}{a} = \frac{S_1S'_1}{b + D}, \quad \text{and } S_1S'_1 = 2D\theta,$$

where $\theta$ is the small angle between the two positions of the mirror.

Therefore

$$x = \frac{2aD\theta}{b + D} = \frac{2aD\omega T}{b + D} = \frac{4a\omega D^2}{v(b + D)},$$

or

$$v = \frac{8\pi naD^2}{x(b + D)},$$

which expresses $v$ in terms of quantities which can be measured.

In the final experiments of Foucault a beam of solar light was reflected horizontally from a heliostat through the aperture S. The sight used was not a fine wire stretched across the aperture, but a microscope scale which consisted of fine lines traced on a piece of silvered glass at a distance of $\frac{1}{10}$ mm. from each other. This scale was placed at S, so that the light passed through it and an image of it was viewed at $a$ in the field of the observing microscope. What was observed therefore was not the displacement of the image of the slit but the displacement of the image of this scale. The revolving mirror was a piece of glass silvered and polished on one face. This was supported in a strong ring frame, and its diameter was 14 mm. The radius of curvature of the fixed mirror M was 4 metres, so that with a single fixed mirror, as in Fig. 221, the distance D in the foregoing formula would be 4 metres. In the actual experiment, however, this was increased to 20 metres by using five fixed mirrors instead of one. For this purpose M was turned a little to one side, so that the light reaching it from the revolving mirror was not reflected back directly to R as already described, but to another fixed mirror of equal radius. From this it was reflected to a third, and then to a fourth, and finally to a fifth, which received it normally, and returned it along its previous path to the revolving mirror, and thence to the field of the observing microscope. The lens L, which had a focal length of 1·9 metres, was placed between the revolving mirror and the first fixed mirror (Fig. 222), and not, as in Fig. 221, between the revolving mirror and the aperture.

The speed.

In order to determine the speed of the revolving mirror, and to control it during an observation, a most ingenious device was resorted to. A finely divided toothed wheel was placed between the observing microscope and the reflecting glass plate, so that the image of its toothed edge appeared in the field of view. This wheel was driven by clockwork at a uniform speed, which would be accurately determined. Now the beam of light entering the field of view is not continuous, but intermittent. It consists, in fact, of a succession of flashes, each flash corresponding to a complete turn of the revolving mirror R. If the beam of light were continuous, the teeth of the revolving disc would be seen rapidly crossing the field at a speed depending only on the rate at which it is driven, and at any considerable speed they could not be distinguished in passing. With the intermittent beam, however, the teeth are illuminated for a very short time once during each revolution of the turning mirror R, and the result is that if the toothed wheel turns through an angle corresponding to any whole number of teeth in the time between two consecutive flashes, the position of the teeth in the field of view will always appear to be the same when they are illuminated, and the teeth will consequently appear to be stationary. If the speed of the toothed wheel be greater or less than this the teeth will appear to have a slow forward or backward motion in the field of view.

The revolving mirror was driven by an air turbine so that its speed could be controlled, and during an observation this was so regulated that the image of the toothed wheel appeared stationary in the field of view. The speed of the toothed wheel being known, the number of teeth passing between two consecutive flashes could be determined, and from this the speed of the rotating mirror was easily found.

The observed displacement of the scale was 0·7 mm., and the final result for the velocity of light was 298,000,000 metres per second.

**304. Discussion of the Revolving Mirror Method.**—In Foucault's investigations the distance D was so small (being only 20 metres with five fixed mirrors) that a large angular deviation of the image was out of the question, and by reason of the many reflections there was necessarily a serious loss of light. In order to obtain a large deflection with a given speed it is necessary to work with a large distance between the two mirrors, and as there is always light lost by reflection and absorption in passing over a great distance it is necessary to attend to the conditions which render the image most brilliant. When the displacement of the image is large the reflecting glass plate is unnecessary and may be dispensed with. The image formed by the returning light can then be observed directly without being weakened by the reflections attending the use of the glass plate. Now the angular deviation of the return image, for a given speed of the revolving mirror, increases with the distance D, and for a given angular deviation the displacement of the image is proportional to the distance between the source and the revolving mirror, or, as it is called, the *radius*. Hence for a large displacement of the image the distance between the mirrors, the radius, and the speed should each be made as large as possible. The second condition is obviously in conflict with the first, for the slit and the fixed mirror must be situated in the conjugate foci of the lens L.

When the lens is placed between the revolving mirror and the slit, as in Fig. 221, the quantity of light returned by M to R varies inversely as the distance D. Thus with a concave mirror of one decimetre diameter placed at a distance of one kilometre the light returned to the revolving mirror would not be as much as $\frac{1}{60000}$ part of the light reflected from it. This quantity is further reduced by atmospheric vibration, diffusion, and absorption, and with large distances it is almost impossible to construct a mirror of such uniform curvature that the rays will fall normally on every part of its surface. In general only a part of its surface will satisfy this condition, so that a portion only of it is effective, and this leads to a further decrease in the brilliancy of the return image. Hence with a concave mirror of one decimetre diameter for each kilometre of distance only a small fraction of the $\frac{1}{60000}$ part of the light reflected from the revolving mirror is really utilised when the apparatus is arranged as in Fig. 221. On the other hand, when the lens is placed between the revolving mirror and the fixed mirror, as in Fig. 222, it is easily seen that if R and M are in conjugate foci of L, then the light reflected from R will fall upon M as long as the axis of the reflected beam falls upon the lens, however great the distance D may be. This arrangement, however, cannot be made, for it is the slit S and not the mirror R that must be in the conjugate focus of M; nevertheless, it may be approximated to by bringing the slit close to the revolving mirror, and the brilliancy of the return image will be increased accordingly—that is, approximately in the ratio of the angular diameter of the lens, subtended at the centre of motion, to that of the mirror M at the same point.

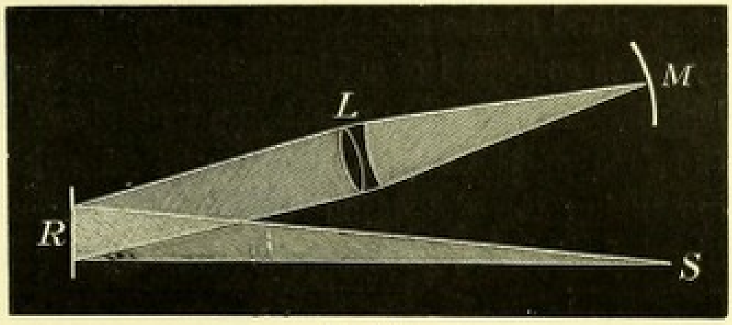


Fig. 222.

The advantages derived from increasing the distance D are attended by serious defects in the return image caused by atmospheric diffusion and vibration. For the light which forms the image, instead of travelling accurately along a definite line ML, between the fixed mirror and the lens, is scattered through a certain angle. This leads to an error in position of some parts of the image, which, when expressed in linear measure, will be proportional to the focal length of the lens. In other words, a limit is soon reached, beyond which the angular accuracy with which a micrometer wire can be set on the image of a star is not increased by increasing the length of the telescope. The result is that the brilliancy of the image cannot be increased many fold by increasing the focal length of the lens without at the same time increasing too much this source of error.

Another possible drawback to the use of the lens in the position shown in Fig. 222 is that an image of it, or of some part of it, will flash through the field of view with every revolution of the mirror, and this will lead to a certain amount of illumination in the field. It is obviously desirable that this, as well as all other extraneous illumination, should be excluded from the field of view, in which a very faint image is to be observed.

**305. Michelson's Experiments.**—The chief objection to Foucault's experiments is that the deflection was too small to be measured with sufficient accuracy, and to remedy this defect Professor Michelson[1] modified the arrangement of the apparatus in such a way that the return image was displaced through 133 mm., or about 200 times that obtained to Foucault. With a large displacement such as this the inclined glass plate could be dispensed with, and the return image was observed directly through a micrometer eyepiece placed on the same stand as the slit through which the light was transmitted. The eyepiece consisted of a single achromatic lens of about 2 inches focal length. In the focus of this, and nearly in the same vertical plane as the face of the slit, a single vertical silk fibre was stretched. In measuring the deflection the eyepiece was so placed that the fibre bisected the slit, and it was then moved till the fibre bisected the deflected image of the slit.[2]

The revolving mirror was a disc of plane glass about 1¼ inch in diameter and 0·2 inch thick. It was silvered on the front surface, so that reflection took place from one surface only. The lens was placed between the two mirrors, as in Fig. 222, so as to secure greater brightness of the return image, and its focal length was 150 feet. The revolving mirror was placed 15 feet inside the principal focus of the lens, and the distance between the two mirrors was about 2000 feet. In the first set of experiments the fixed mirror was plane, being about 7 inches in diameter, and a small telescope was attached to it (with the line of collimation at right angles to the surface of the mirror) for the purposes of adjustment.[3] The lens was 8 inches in diameter and was not achromatic, but on account of its great focal length, as compared with its aperture, the want of achromatism was not appreciable. The "radius," or the distance between the slit and the revolving mirror, was about 28 feet. In making an experiment it was found necessary to incline the axis of the revolving mirror slightly to the right or left, so that the light falling directly on it from the slit

[1] A. A. Michelson, *Astronomical Papers for the American Ephemeris and Nautical Almanac*, vol. i. part iii. p. 117, 1880.

[2] It may be observed that in the experiments of Michelson and Newcomb the image of the slit was worked with rather than the image of a wire, or scale, placed across the slit, as in the experiments of Foucault. This was necessary on account of the great distance between the mirrors, for an accurate image of a line could not be formed when the distance was considerable. Further, when the distance is large the transmission through the atmosphere renders the image very unsteady, especially about the middle of the day. It was only during the hour after sunrise, or the hour before sunset, that a sufficiently steady image of the slit could be obtained in these experiments.

[3] The particulars of adjustment are described in Professor Michelson's paper, *loc. cit.*

should not be reflected into the eyepiece but should pass either above or below it. Without this precaution the light of the slit would be flashed into the field of view at every revolution of the mirror and would overpower that of the image to be observed.

The revolving mirror was driven by an air turbine controlled by a cord leading from its valve to the observer's table. To measure the speed of rotation a tuning-fork, bearing on one prong a steel mirror, was used. This was kept in vibration by an electric current from five "gravity" cells. The fork was so placed that the light from the revolving mirror fell upon it and was reflected to a piece of plane glass (placed in front of the eyepiece of the micrometer), inclined at 45°, and thence to the eye. When the fork and the revolving mirror are both at rest, an image of the revolving mirror is seen. When the fork vibrates this image is drawn out into a band of light. When the mirror revolves this band breaks up into a number of moving images of the mirror, and finally, when the mirror makes as many turns as the fork makes vibrations, these images are reduced to one, which is stationary. This is also the case when the number of turns is a submultiple of the number of vibrations. When it is a multiple (or simple ratio), the only difference is that there are more images. Hence to make the mirror execute a certain number of turns, it is simply necessary to pull the cord attached to the valve to the right or to the left, until the image of the revolving mirror comes to rest. In this way it was possible to keep the mirror at a constant speed for three or four seconds at a time, and this was sufficient for an observation. In a large number of the experiments the speed was approximately 258 revolutions per second, and the mean result for the velocity of light in vacuo was put down at

$$v = 299{,}910 \pm 50 \text{ kilometres per second.}$$

In a subsequent series of supplementary measures Professor Michelson[1] determined the velocity of white and coloured light in air, water, and bisulphide of carbon. In these experiments the arrangement of the apparatus was the same as before; the fixed mirror, however, was slightly concave, and had a diameter of 15 inches. Particular attention was paid to the appearance of the return image, in order to detect if it indicated any difference in the velocities of the different colours in air, such as was supposed to have been observed by Young and Forbes. The actual width of the slit was 0·19 mm. and the width of the return image was only 0·25 mm. The colour of

[1] A. A. Michelson, *Astronomical Papers for the American Ephemeris and Nautical Almanac*, vol. ii. part iv. p. 237, 1885.

the central portion of this image was always yellowish, and occasionally both borders were observed to have a pale violet tinge. There was consequently no indication of a spectral drawing out of the image such as would result if the different colours travelled with different velocities. A difference of velocity such as that obtained by Young and Forbes should have yielded a spectral image of about 10 mm. in width. Finally, experiments were made in which a plate of red glass covered one half of the slit so that one half of the return image was white while the other was red. The two halves of the image were found to be exactly in line, and showed no break or displacement such as would attend a difference of velocity in the different colours.

The weighted mean of 318 observations on sunlight gave for the velocity in a vacuum

$$v = 299{,}865 \text{ kilometres},$$

while the weighted mean of 267 observations on electric light gave

$$v = 299{,}835 \text{ kilometres}.$$

**306. Newcomb's Experiments.**—The most recent investigation of the velocity of light by the revolving mirror method was made at Washington in the years 1880-82 by Professor Simon Newcomb,[1] and this determination is perhaps the most reliable and complete that has yet been recorded.

In these experiments the quantity directly measured was not the linear displacement of the return image but its angular deviation. The plan of the apparatus is shown in Fig. 223, and it will be seen that the method followed by Foucault and Michelson of placing the lens between the two mirrors was not adopted by Newcomb. After due consideration the method of Fig. 221 was employed, the lens being placed between the slit and the revolving mirror. The general feature of the apparatus is that two telescopes were disposed with their axes at right angles to each other, one, F, termed the sending telescope, being used to cast a pencil of light on the revolving mirror, and the other, L, the observing or receiving telescope, being employed to receive the return beam.

The light of the sun, thrown from a heliostat, entered the slit S of the sending telescope, and after passing along the tube F was reflected by a plane mirror at the elbow C through the object-glass J. It then fell upon the revolving mirror contained in the box *m*, and was there reflected along the line Z to the distant fixed mirror. The object-glass of the receiving telescope was immediately below J, and its tube L

[1] S. Newcomb, *Astronomical Papers for the American Ephemeris and Nautical Almanac*, vol. iii. part iii. p. 113, 1885.

was mounted in adjustable Y's on a frame NN, which moved horizontally around a vertical axis coinciding with the axis of rotation of the mirror. The farther end of this telescope was fitted with a pair of microscopes, $p$ and $h$, for reading the divisions of the graduated arc below.

With the apparatus disposed in this manner it was necessary to elongate the revolving mirror to such an extent that the light, when

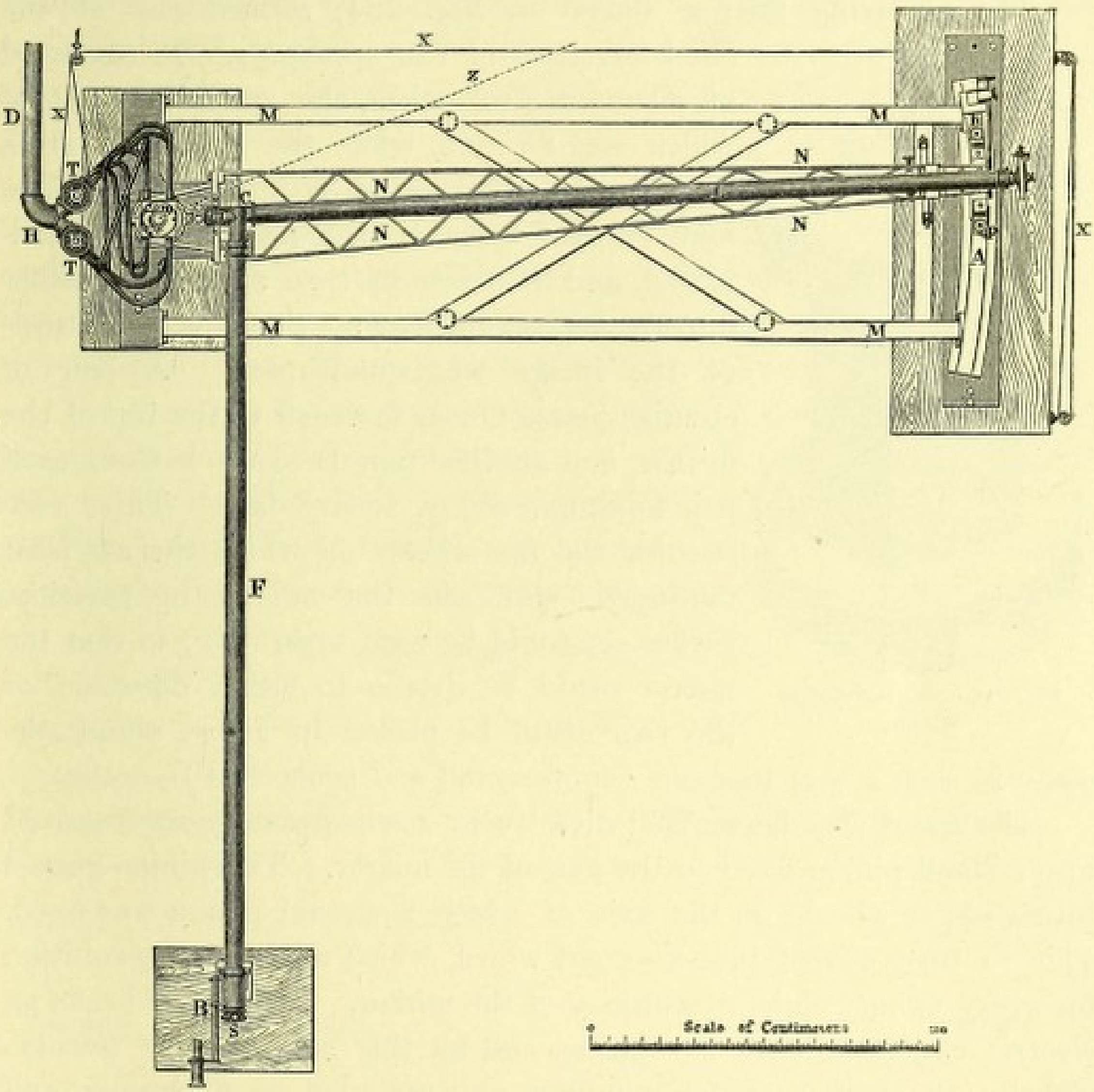



Fig. 223.

cast upon the upper part of it by the upper or sending telescope F, should, after reflection at the fixed mirror, return, not to the same part of the revolving mirror, but to a part somewhat lower down, so as to enter the lower or observing telescope L. By this means the light, irregularly reflected at the surface of the revolving mirror where the incident beam falls upon it, is prevented entering the observing telescope, and greater darkness of the field is thereby secured.

With this arrangement almost all the extraneous light can be shut out of the field, and a very faint image of the slit can be observed. Further, by causing the mirror to revolve first in one direction, and then in the other, deviations on opposite sides of the zero can be observed, and in this manner the zero error can be eliminated.

In order to strengthen the illumination of the return image, and to avoid loss from any slight displacement of a single mirror, two fixed mirrors were used side by side. These were concave, and each had an aperture of about 40 cm., and a radius of curvature of about 3000 metres.

The revolving mirror, shown in Fig. 224, formed one of the chief novelties of the apparatus. It consisted of a square steel prism, the vertical height of which was 85 mm., while the horizontal cross section was a square of 37·5 mm. side. The four vertical faces of the prism were nickel-plated, and each face in turn acted as reflector during the revolution, so that the brightness of the image was quadrupled. A pair of circular plates C was fastened to the top of the mirror, and another pair D to the bottom, each pair holding a set of twelve fans. These constituted the fan wheels on which the air blast impinged, and set the mirror in rotation. Either set could be used separately, so that the mirror could be driven in either direction, or the two could be placed in action simultaneously in such a way that one counteracted and controlled the other.

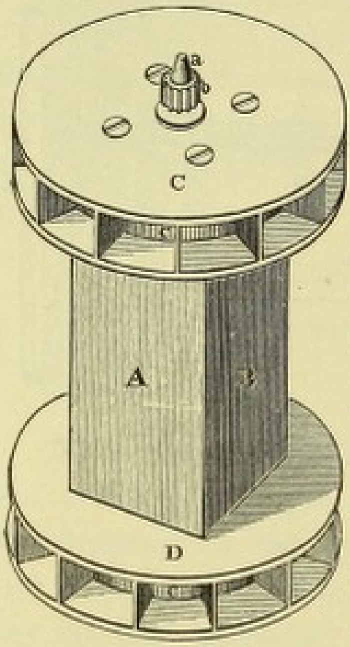


Fig. 224.—The Revolving Mirror.

The speed was determined directly by a wheelwork system geared into a small pinion fixed to the axis of the mirror. This pinion geared into a larger wheel,[1] on the axis of which a second pinion was fixed. This in turn geared into a second wheel, which made one revolution for every twenty-eight revolutions of the mirror. This wheel broke an electric circuit at each revolution, and by this means every twenty-eighth turn of the revolving mirror was recorded on a chronograph along with the beats of a sidereal break circuit chronometer.

In making an experiment the observing telescope was first set in a fixed position corresponding to some desirable deflection of the return image, and the speed of the revolving mirror was then so adjusted that the return image entered the field of view and came to rest upon the micrometer wires of the eyepiece. In order that the observer at the eyepiece should be able to regulate the speed of the mirror, one of the valves T could be controlled by means of an end-

[1] It was found that the best metal wheels were worn out almost at once under the high speeds at which the first wheel was forced to move. For this reason it was necessary to resort to raw-hide as the material for the first wheel, and this proved successful.

less cord X, which passed through pulleys around the side of the instrument to the front of the observer's table. This valve being shut the other was opened and the mirror set in motion. When the speed reached the proper limit the image was seen entering the field, and as it approached the cross wires the chronograph was started. The cord X was then moved so as to slightly open the valve to which it was attached. This allowed a slight counterblast to act upon the other fan wheel, and by this means the speed could be regulated with the greatest delicacy. When the image was steadily adjusted on the wires it was kept there for about two minutes, and the corresponding record on the chronograph furnished the speed of rotation. It was remarked that the higher the speed the greater the delicacy with which the image could be adjusted to the cross wires. The observing telescope was then set on some division on the other side of the zero, and the mirror was made to rotate in the opposite direction, and a run taken as before.

The final conclusion from these experiments was that the velocity of light in air is 299,728 kilometres per second, and in vacuo 299,810 km., with a probable error estimated at 40 or 50 km.

Using only the results of those determinations which were supposed to be free from any constant error, the concluded velocity was

In vacuo, $v = 299,860 \pm 30$ kilometres.

Professor Newcomb gives the following table :—

RESULTS OBTAINED FOR THE VELOCITY OF LIGHT IN VACUO

| Observer. | Velocity in Kilometres. |
|---|---|
| Foucault, at Paris, in 1862 . . . . . . . . | 298000 |
| Cornu, ,, 1874 . . . . . . . . . | 298500 |
| ,, ,, 1878 . . . . . . . . . | 300400 |
| Foregoing, as discussed by Listing . . . . . | 299990 |
| Young and Forbes, 1880-81 . . . . . . . . | 301382 |
| Michelson, at Naval Academy, 1879 . . . . . . | 299910 |
| ,, Cleveland, 1882 . . . . . . . | 299853 |
| Newcomb, at Washington, 1882— | |
| (*a*) Using only results supposed free from constant error . | 299860 |
| (*b*) Including all observations . . . . . . . | 299810 |

**307. Comparison of the Velocities in Different Media.**—The first application of the revolving mirror method was not directed so much to a determination of the absolute velocity of light in air or any other medium as to the comparison of the velocities in different transparent media. The emission theory demanded a greater velocity for light in the more highly refracting media, while the wave theory required the reverse, and the burning problem of the time was to obtain a direct experimental test. For this reason the earlier experiments of Foucault were directed to determine whether the velocity of light in air is greater or less than in water. If light travels slower in water than in air, then it is clear that if a tube of water be placed between the

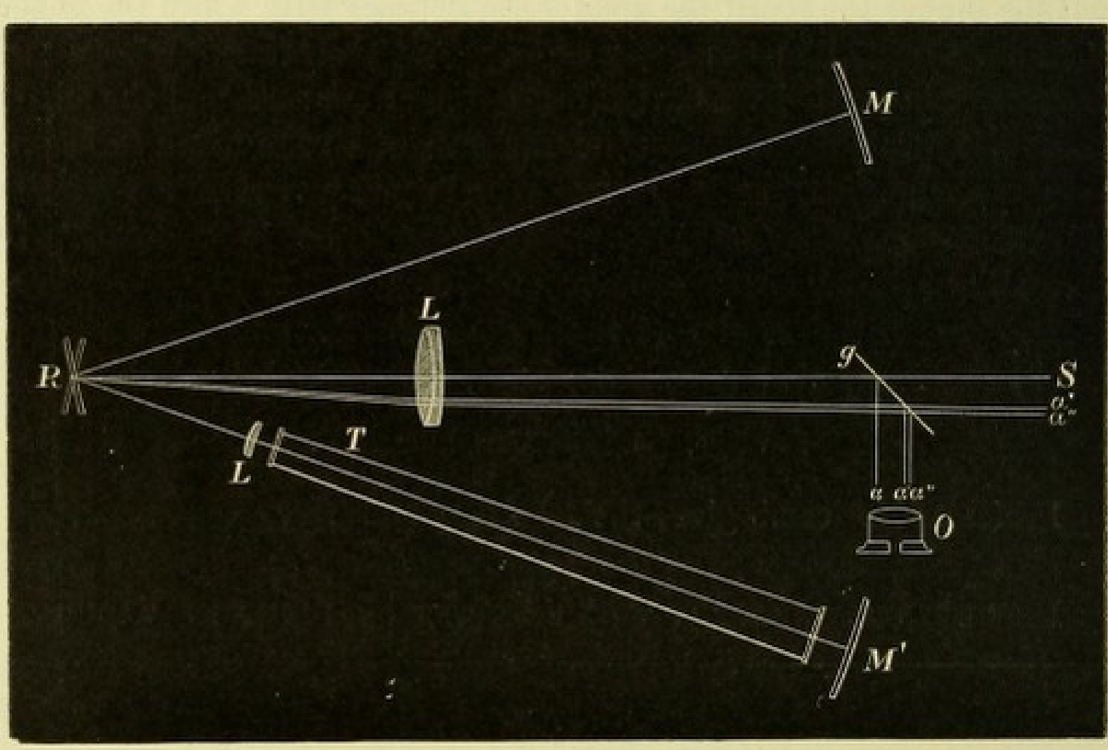



Fig. 225.

revolving mirror and the fixed mirror, so as to be traversed by the reflected beam of light, then a longer time will be spent in the double journey between the mirrors, and the deflection of the return image will be increased. On the other hand, the deflection will be diminished when the water is interposed if the velocity in it is greater than in air. To test this point, therefore, it is only necessary to interpose a column of water and observe how much, and in what direction the deflection is changed at any given speed of the mirror.

The difficulty of comparing the two deflections at the same speed is eliminated by arranging the experiment so that the two return images are seen in the field of view simultaneously. This was the plan adopted by Foucault, and was effected by placing the tube of water T (Fig. 225) between the revolving mirror R and a second fixed mirror M′ of the same curvature as M, and having its centre at R. When the revolving mirror is set in motion, a return image will appear as before in the field of view by reflection from M. This may be termed the air image. For the same reason there is another return image arising from the light cast upon M′ and reflected there. This may be called the water image. For small speeds these two images will be superposed at $a$, but when the speed is increased they separate, the water image $a''$, being more displaced than the air image $a'$, showing that light travels faster in air than in water.

The image observed was that of a fine vertical wire stretched across the slit parallel to its length, and in order that the two images should be simultaneously in focus, it was necessary to place a slightly converging lens L′ in front of the tube T. This correction was necessary because the lens L is so placed that S and M are in conjugate foci when the space between the mirrors is filled with air, consequently S and M′ cannot be in conjugate foci of L when the water tube is interposed.

In this manner Foucault proved that light travels faster in air than in water, but he made no estimate of the ratio of the velocities. This was done in 1883 by Professor A. A. Michelson,[1] the apparatus being disposed in the manner devised by Foucault. The tube T contained distilled water and was about 3 metres long, the distance between the mirrors was about 5 metres, and the "radius" 10 metres. The speed was 256 revolutions per second, and the ratio of the velocity of light in air to that in water was found to be 1·330. The refractive index of water for yellow light is 1·333, and the difference between these two numbers was considered within the limits of experimental error.

In the case of carbon bisulphide the refractive index for the mean yellow rays is 1·64, and the ratio of the velocities was found to be 1·758. This result is about 7 per cent too high, and the discrepancy could not be regarded as within the limits of experimental error (see Art. 310).

Experiments were also made by Professor Michelson on the velocities of differently coloured lights in carbon bisulphide. For this purpose the solar light was passed through a direct vision spectroscope before it fell upon the slit, and by turning the prism through a small angle, either end of the spectrum could be observed. The results indicate that the orange-red light travels 1 or 2 per cent faster in carbon bisulphide than the greenish-blue light.

[1] A. A. Michelson, *loc. cit.*

## Wave Velocity and Group Velocity

**308. The Velocity determined by the Methods of Römer and Fizeau.**—Shortly after the announcement of the results of Messrs. Young and Forbes, Lord Rayleigh[1] raised and discussed the question as to what it is that is really determined in observations on the velocity of light such as we have described. If we could deal with a single wave and observe its progress, we could determine its speed—that is, the wave velocity. In the case of light, however, we cannot follow the motion of a single wave. Here we deal with a group of waves, and measure the velocity with which some impressed peculiarity travels. Thus in determinations by the eclipses of Jupiter's satellites, or by Fizeau's method, the light is rendered intermittent, and the velocity observed is that of a limited train of waves—that is, the group velocity.

The relation of the group velocity $u$ to the wave velocity $v$ will be found in a note appended to chap ii. (p. 57). If $k=2\pi/\lambda$ we have

$$u=\frac{d(kv)}{dk},$$

so that $u$ is identical with $v$ only when $v$ is independent of $k$—that is, of the wave length, and this relation is generally supposed to hold for light traversing free space. The equation shows that a complete knowledge of $v$ completely determines $u$, but a complete knowledge of $u$ does not determine $v$ without the aid of some auxiliary assumption. The usual assumption is that $v$ is independent of the wave length, in which case $u$ will also be independent of the wave length. If, however, $v$ is not independent of $\lambda$, we may substitute some dispersion formula,[2] such as

$$v=\mathrm{A}+\mathrm{B}k^2+\mathrm{C}k^4+\ .\ .\ .$$

Using the formula $v=\mathrm{A}+\mathrm{B}k^2$, and taking the wave lengths of the orange-red and green-blue lights to be in the ratio 6 : 5, Lord Rayleigh finds that the wave velocity would be nearly 3 per cent less than the

[1] Lord Rayleigh, *Nature*, 25th August 1881 and 17th November 1881.

[2] [If in vacuo $v$ is independent of $\lambda$, and if the refractive index of air is given by Kayser and Runge's formula $\mu=a+bk^2+ck^4$ (p. 137), we have

$$u=kv\frac{d}{dk}\log kv=kv\frac{d}{dk}\log\frac{k}{\mu}$$

or

$$u=v\left(1-\frac{2bk^2+4ck^4}{a+bk^2+ck^4}\right)=v\left\{1-\frac{2b}{a}k^2-\left(\frac{4c}{a}-\frac{2b^2}{a^2}\right)k^4\right\}$$

as the relation between the group and wave velocity in air. Inserting the numerical values from the page cited, we find for $\lambda=500$ millionths of a millimetre $u=v(1-0{\cdot}000013)$. Thus the difference in air would be quite inappreciable.]

velocity determined by the eclipse method, the relation between $u$ and $v$ being

$$v = u(1 - 0 \cdot 0273).$$

**309. The Velocity determined by the Aberration Method.**—In measurements depending on the aberration of light the velocity determined does not depend upon the observation of the rate of propagation of any impressed peculiarity. It is consequently not the group velocity $u$, and, according to the usual theory of aberration, must be the wave velocity $v$. We have therefore no reason *a priori* to expect that the velocity given by Bradley's method should be the same as that found by the methods of Römer and Fizeau, unless the wave velocity be identical with the group velocity, which postulates that the speed is independent of the wave length. The close agreement of the velocity $v$ found by observation of the coefficient of aberration, and that ($u$) given by Fizeau's method leaves little room for the supposition that $u$ is different from $v$ in air, and points to the conclusion that the velocity is independent of the wave length, or that all colours traverse interstellar spaces with the same speed.

**310. The Velocity determined by the Revolving Mirror Method.**—In this method of determining the velocity of light, the first reflection of the beam takes place at the surface of a revolving mirror, and on account of the motion of this mirror, the angle of incidence varies, while the direction of the incident beam remains constant. It follows, therefore, that the successive waves are thrown off at different angles, and consequently the reflected beam is a curved stream of light, in which the successive wave fronts are inclined to each other at an angle depending upon the speed of the mirror.

This reflected stream as it travels through space sweeps past the fixed mirror, and a segment of it is reflected there. By properly adjusting this mirror the reflected segment is returned in such a direction that it falls upon the revolving mirror and enters the eyepiece of the observing microscope. An image is thus depicted in the field of view, and its illumination arises, not from a continuous stream of light, but from a succession of flashes, each flash being produced by a reflected segment, and corresponding to a revolution of the moving mirror.[1] The image being produced by a succession of flashes, it would appear at first sight that the velocity determined by this method ought to be the same as in Fizeau's experiment, namely, the group velocity $u$. But Lord Rayleigh[2] has remarked that as the successive

[1] It is to be noticed that the stream leaving the revolving mirror is itself intermittent.

[2] Lord Rayleigh, *Nature*, vol. xxv. p. 52, 17th November 1881.

wave fronts are inclined to each other after reflection from the moving mirror, then if the velocity depends upon the wave length there will be a rotation of the wave fronts in the air during the transit between the two mirrors, so that the velocity determined by this method might be some function[1] of $v$ and $u$.

It has been pointed out, however, by Professor J. Willard Gibbs[2] that although the individual waves rotate, yet the wave normal to the group remains unchanged; or, in other words, if we fix our attention on a point moving with the group, and therefore with velocity $u$, the successive waves all pass through that point with the same orientation. He concludes, therefore, that the velocity determined by this method is the group velocity $u$.

Treating Michelson's experiments on bisulphide of carbon on this supposition he finds for the mean of the D and E lines

$$\frac{V}{u} = 1{\cdot}745,$$

where V is the vacuum velocity. The experimental result obtained by Michelson was 1·76 for light in which the maximum brilliancy was between D and E, but nearer to D than to E. The agreement is thus tolerably exact, and would be made more so by taking a value nearer D instead of the mean of D and E.

## Relative Motion of Matter and the Ether

**311. Effects of Relative Motion.**—When the observer is not at rest, but is moving through the medium in which a system of waves is being propagated, or when the source itself is moving through the medium, or more generally, when the source, medium, and observer are moving relatively to each other, certain changes may occur in the observed effects. These we shall now briefly notice.

It is known, for example, that when the observer is moving through the air towards a source of sound, or when the source is moving towards the observer, the frequency (or pitch) of the note appears to the observer to be increased, whereas, when they are moving away from each other, the reverse occurs. Thus one effect of relative motion is a change in the observed frequency. This is the well-known Doppler effect, and if light be a wave motion in a medium, something of the same sort should occur when the source and observer are moving relatively to each other. This change of frequency has been already

[1] The value deduced by Lord Rayleigh was $v^2/u$, and the value found by Professor A. Schuster was $\frac{v^2}{2v-u}$ (*Nature*, vol. xxxiv. p. 439, 1886).

[2] J. Willard Gibbs, *Nature*, vol. xxxiv. p. 582, 1886.

noticed and calculated in Art. 286, and may be examined spectroscopically.

Again, if the medium itself be in motion, the waves will drift in the direction of motion, so that the velocity with which they travel will be greater in the direction in which the medium is moving than in the opposite direction. A motion of the medium will thus cause a change in the velocity of propagation. Further, when there is motion of the source, medium, or observer, the intensity may be different in different directions, but this may be compensated in some cases by a change in the radiating power of the source in different directions imposed by the motion of the source.

Finally, when the observer is moving relatively to the source a change may occur in the direction in which the waves appear to travel. This change in direction is known as aberration, and on it is founded the method by which Bradley was led to an estimation of the velocity of light (Art. 18).

Hence if light be a wave motion, and if its mode of propagation in the ether be at all similar to that of waves in an ordinary material medium, then relative motion of the source, medium, and observer should produce effects similar to those mentioned above, namely, changes in frequency, velocity, intensity, and direction. We should bear in mind, however, that in reasoning on this subject we approach it with ideas derived from the study of material media, and that in drawing our conclusions we are merely arguing by analogy from phenomena which occur in media which we can examine and control, to those of a medium of which we are almost wholly ignorant. In addition we have no knowledge as to the real nature of the periodic change which occurs in what we term "the vibration," whether it be a motion, as ordinarily conceived, or a periodic change of some property or condition of the ether. We do know, however, that the velocity of propagation of light is greater in free space than when passing through a region in which matter is diffused, and that the velocity in a space occupied by matter is different for waves of different lengths. As to whether the ether itself is really modified by the presence of the matter (for example, if there is a real change of density as Fresnel supposed), or as to whether the change is only virtual, being merely an affection arising from the influence of the matter on the velocity of propagation of waves through the ether in a space in which matter molecules are diffused, is still a subject of speculation. The prevailing conception seems to have been that the ether within a piece of matter really differs in some quality from the ether in free space. If there is a real change of density in the ordinary sense of the term, then the doctrine of an incompressible ether would present a serious difficulty.

It is sufficient, however, to regard the modifications imposed by matter to be merely virtual, so that the effect of the molecules of matter distributed through a region of the ether is to influence the propagation of waves through that region in a manner similar to that which would be produced by a real change in those properties of the medium which determine wave propagation through it. Thus if a system of floats or other bodies be suspended in a certain region of a fluid, the propagation of waves through this region will differ from that in the free regions, and this amounts to a virtual change in the elasticity or density of the fluid in that region. Thus the fluid itself may be really unaltered by the presence of foreign bodies distributed through it, but it may be virtually altered as far as wave propagation is concerned.

The fact before us at present is that the rate of propagation of waves in the ether is modified by the presence of matter, and the question which at once arises is as to whether the velocity of propagation of ether waves in a region occupied by matter is influenced by the motion of the matter through space. This question we shall now consider in conjunction with the phenomenon of aberration.

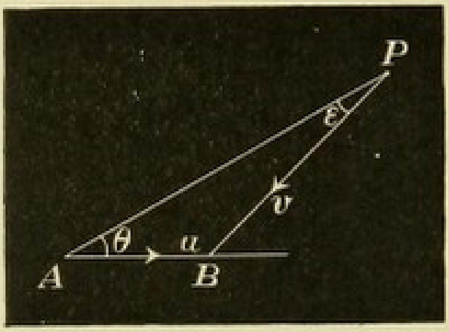



Fig. 226.

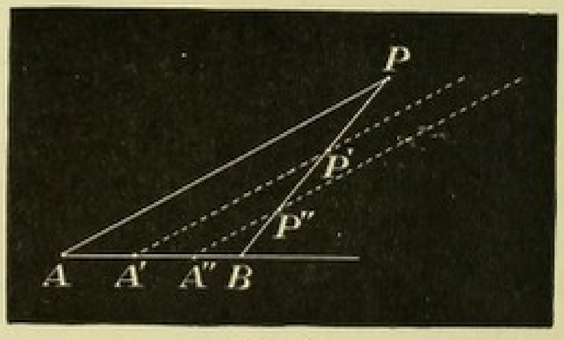



Fig. 227.

**312. Aberration.**—In order to illustrate the principles of aberration let us take the case of an observer A moving in the direction AB (Fig. 226) with a velocity $u$, while a particle P moves with a velocity $v$ in the direction PB, and let the lengths AB and PB represent the magnitudes of $u$ and $v$ respectively, then it is clear that A and P will reach B at the same instant and will collide there. In other words, the direction in which P must be fired in order to strike A is not along the line PA but along a line PB, making a certain angle with PA. This is illustrated by the well-known fact that in order to hit a bird flying across the line of fire the gun must be aimed, not directly at the bird, but at a point somewhat in advance of it. On the other hand, regarding the subject from the point of view of the observer A, it is clear that as P moves along PB, and A moves along AB, the line joining P and A will be always parallel to its initial direction AP. Hence what A observes is that P approaches him in a direction parallel to AP and finally strikes him. The direction in which P appears to approach A is consequently parallel to AP, while the direction in which it would appear to move if A were at rest would be PB. Thus the motion of the observer alters the apparent direction of motion of the particle P, and the angle APB between the apparent and real directions of motion is termed the aberration.

In order to obtain a clearer view of the matter let us suppose that a straight tube is laid from A to P, and let this tube be carried by A, so that the direction of its axis remains fixed—that is, so that it moves with a velocity of translation $u$ in the direction AB (Fig. 227). Then it is clear that when A occupies the positions A′, A″, etc., P will occupy positions P′, P″, etc., such that the lines A′P′, A″P″, are parallel to AP—that is, if the observer looks along the tube he will always see P on its axis; or, in other words, P will appear to move towards him along the axis of the tube. This, then, is the apparent direction of motion of the particle.

An expression for the aberration $\epsilon$ arising from the motion of the observer may now be written down at once. For if the angle PAB be denoted by $\theta$, we have

$$\frac{u}{v}=\frac{\text{AB}}{\text{PB}}=\frac{\sin\epsilon}{\sin\theta},$$

therefore

$$\sin\epsilon=\frac{u}{v}\sin\theta.$$

The quantity $u/v$ is termed the aberration constant, and it is clearly the sine of the maximum angle of aberration, or the tangent of the aberration, when the directions of motion are at right angles to each other.

Instead of regarding P as a moving particle, we might regard it as a certain definite or "labelled" element of a wave front advancing in the direction PB; and further, if we think of the tube AP as a telescope, then we see that the direction in which a telescope must be pointed in order to receive light advancing in the direction PB, will be BP when the observer is at rest, but will be AP when the observer is moving. Hence when the motion of the earth around the sun is taken into account, it should follow that the direction in which a star is seen should be different at different times of the year. In other words, the motion of the earth should cause the stars to appear to describe small orbits around their true positions, these orbits being small ellipses on the celestial sphere—stars on the ecliptic describing right lines, and stars at the pole of the ecliptic describing small circles. This is the phenomenon observed by Bradley, and the explanation put forward is that it arises from the orbital motion of the earth compounded with the velocity of light.

The general explanation of the aberration of the stars in this manner is so simple, and the value of the velocity of light deduced from it is so close to that obtained by direct experiment (and from observations on the eclipses of Jupiter's satellites), that the truth of the explanation can scarcely be doubted. On the emission theory, indeed, this explanation would appear to be particularly acceptable, but, when the subject is examined a little more closely, the explanation is not so simple as it appears at first sight. This arises from the possible drift which may occur when a wave (or a light corpuscle) is being propagated through moving matter.

**313. The Drift produced by Moving Matter.**—In Fig. 227 we have considered AP as the axis of a tube, and this tube has been supposed empty, so that the velocity of P remains the same whether we suppose it moving within the tube or outside it. If, however, the tube be filled with a medium in which P moves with a velocity $v'$, and if we suppose the motion of P to be unaffected by the motion of the medium filling the tube, then clearly the inclination of the tube must be altered to suit the new velocity $v'$. In other words, the aberration will depend on the nature of the substance filling the tube.

In drawing this conclusion, however, we have left out of account an important consideration, namely, that on account of the motion of the substance filling the tube, P may be dragged in the direction in which the tube is moving, and that by reason of this drag (or drift) a compensation may occur, so that the inclination of the tube (or the observed aberration) may remain unaltered. Further, it must also be taken into account that the direction of motion of P may change, or refraction may occur at the surface of the medium within the tube, so that we cannot say *a priori* whether the aberration of a star should or should not be the same when the tube is filled with a given substance as when it is empty.

As a matter of fact the experiment has been tried by Airy and Hoek, and the result is that the aberration of the fixed stars is observed to be the same whether the telescope be filled with water or air. The problem, therefore, before us at present is to account for this fact on the supposition that the light waves drift with the matter through which they are moving, and to determine the law of drift so that compensation may occur, and the aberration remain independent of the medium filling the telescope.

Let AP be the direction of the axis of the telescope and P′P the direction of the incident light in space. Then if the telescope were merely an empty tube the element of wave front at P would move to B′ in the original direction P′P. But when the telescope is filled with a refracting medium the axis AP is normal to the refracting surface, and therefore the angle of incidence is APB′, so that the direction of the refracted ray would be a line PQ, making an angle $r$ with AP when the telescope is supposed stationary.

Now let P′B′ represent V the velocity of light in a vacuum, and let A′B′ represent $u$ the velocity of the observer; then the direction in which the tube must be pointed when empty is A′P′. Hence if the direction of the telescope is to remain unchanged when filled with a refracting substance, we must have AP parallel to A′P′ in Fig. 228. Consequently, if this figure be constructed so that PQ represents (on the same scale) the velocity of light in the refracting substance at rest, then, if there is no drift, P will proceed through the tube as if the medium occupying it were at rest, and will consequently not reach Q until A has reached a point B determined by the equality AB $= u =$ A′B′. But if the waves drift, P will be dragged with the medium so as not to travel along PQ, but along some path PB in advance of PQ. When PB is such that AB $= u$, then the drift is such that complete compensation has taken place. To express this we may call QB the velocity of drift—that is, the rate at which P is dragged in the direction of motion. Denoting it by $u'$ we have

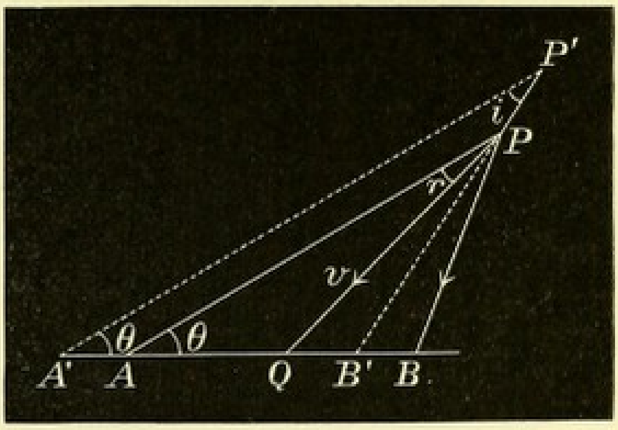


Fig. 228.

$$\mathrm{AQ} = u - u',$$

and consequently

$$\frac{u - u'}{v} = \frac{\mathrm{AQ}}{\mathrm{PQ}} = \frac{\sin r}{\sin \theta} \qquad (1).$$

But denoting A′P′B′ = APB′ by $i$ we have

$$\frac{\sin \theta}{\sin i} = \frac{\mathrm{V}}{u} \qquad (2),$$

where V is the vacuum velocity. Combining these equations we have at once

$$\frac{u - u'}{v} = \frac{\sin r}{\sin i}\frac{u}{\mathrm{V}} = \frac{u}{\mu \mathrm{V}}$$

that is,

$$u - u' = \frac{u}{\mu}\frac{v}{V} = \frac{u}{\mu^2}$$

or finally,

$$u' = \left(1 - \frac{1}{\mu^2}\right)u.$$

The law of drift consequently is that the ether waves must be carried by the moving matter with a velocity $u'$ in the direction of motion, and this velocity is less than the full velocity of the matter in the ratio $(\mu^2 - 1)/\mu^2$.

This is the law deduced by Fresnel on the supposition that the "ether density" is different in different substances, and that the velocity of propagation of light in any substance varies inversely as the square root of the ether density (Art. 207). Thus if the ether density in free space be denoted by $\rho$, while that in a given piece of matter is $\rho'$, then as this piece of matter moves through space it may be regarded as carrying its contained ether with it as if fixed to it, while the external ether is dashed away in front and streams round behind, after the manner of a fluid in which a solid body is moving. Or, on the other hand, we may regard the piece of matter as moving through the ether like a network through a liquid, so that to a person moving with the matter the ether would appear to flow in at the front of the body and out at the rear, the total quantity of ether within the body remaining constant. From this point of view the ether outside the body may be regarded as fixed, while each unit of volume of the body carries with it as permanently attached to it a quantity $\rho' - \rho$ with the full velocity $u$ of the body. This is equivalent to saying that the whole ether within the body is not carried forward with the velocity of the body, but with a velocity $u'$ less than $u$, and determined by the equation.[1]

$$\rho' u' = (\rho' - \rho)u.$$

The equation for $u'$ is therefore

$$u' = \left(1 - \frac{\rho}{\rho'}\right)u = \left(1 - \frac{1}{\mu^2}\right)u.$$

Hence if the ether inside the matter is moving with this velocity ($u'$) relatively to that outside, and if the light waves in traversing this moving ether are carried with it with the full velocity $u'$, then the drift will be such as would render the angle of aberration independent of the substance with which the telescope is filled.

[1] This equation may also be obtained by considering the flux per unit area in the front of the body as $\rho u$, while that within the body is $x\rho'$, therefore $x = u\rho/\rho'$. Hence the velocity of the ether within the body relatively to that outside is

$$u' = u - x = (1 - \rho/\rho')u.$$

Whether the drift of the waves is due to a motion of the ether caused by the moving matter—that is, by the ether being dragged with the matter—or whether, on the other hand, the ether is to be regarded as remaining stationary everywhere, while the matter moves through it, so that the wave drift is caused in some way by the action of the moving matter on the waves as they are passing through it, is a speculation on which it is well to keep an open mind. With regard to this we shall now consider some of the experiments by which it has been shown that light waves drift when passing through moving matter. It is to be remarked, however, that these experiments have always been put forward as proving that the ether itself is carried with the moving matter, and that the wave drift is the result of the motion of the ether.

**314. Fizeau's Experiment.**—The law of drift contained in Fresnel's formula has been directly verified by M. Fizeau[1] in a celebrated experiment in which the moving substance was water. A pencil of light from a narrow slit S (Fig. 229) falls upon a plate of parallel glass

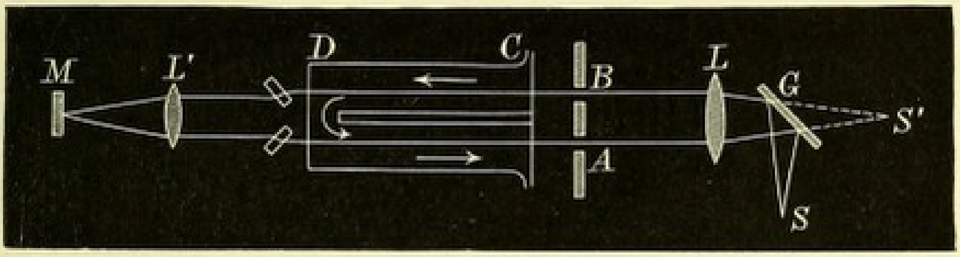



Fig. 229.

G, from which it is reflected and passes through a lens L, from which it emerges in a parallel beam and passes through two apertures A and B. The vessel CD is divided into two chambers by a partition E, and a current of water can be forced through it entering one chamber and leaving the other as indicated by the arrows. The ends of the vessel are closed with plates of parallel glass, and the light from the apertures A and B, after passing through the vessel, is received by a lens L′ and focussed on a mirror M. After reflection from M it again traverses the vessel CD, passes through the apertures, and comes to a focus at S′. It is to be noticed that the light when it enters the vessel through A leaves it through B and *vice versâ*, so that the two pencils which are focussed at S′ have each traversed similar paths, one from A to M and back from M to B, while the other passes from B to M and returns from M to A. If the paths traversed by these beams differ slightly, interference will take place and the image S′ will be crossed by a system of fringes. Now if a current of water be forced through CD, the pencil which enters at B and returns through A will travel with

[1] Fizeau, *Ann. de Chimie et de Physique*, third series, tom. lvii. p. 385, 1859.

the current and the other pencil will travel against the current, so that if the motion of the water has any effect on the rate of propagation of the light, the time of passage of one pencil will differ from that of the other, with the result that an extra phase difference will exist at S′, and the interference fringes will be displaced by a corresponding amount.

Let V be the velocity of light in vacuo, $v$ the velocity in water, and $u$ the velocity of the water, and let the velocity with which the waves are carried by the water be $ux$, then the velocity of the ray which travels with the current will be $v+ux$ and the velocity against the current will be $v-ux$, so that the difference of the times required by the two pencils will be

$$\frac{l}{v-ux}-\frac{l}{v+ux},$$

where $l$ is the length of the water path, that is 2CD. This determines the phase difference introduced by the motion of the water, and the corresponding displacement of the fringes at S′ when measured gives the value of $x$.

M. Fizeau obtained a sensible displacement of the fringes when the velocity of the water was 2 metres per second. With a velocity of 7 metres the effect was measurable, and the result supported the formula proposed by Fresnel, viz. that ether waves travelling in the interior of transparent media are carried forward, but with a velocity less than the velocity of the medium in the ratio $(\mu^2-1)/\mu^2$, where $\mu$ is the refractive index.

This result has been further confirmed by the more recent work of Messrs. Michelson and Morley.[1]

**315. Experiments of Michelson and Morley.**—The subject of the *relative* motion of the earth and the luminiferous ether has also been examined experimentally by Messrs. Michelson and Morley.[2] Their conclusion is that if there be any relative motion between the earth and the adjacent ether it must be small. Thus the ether in the neighbourhood of the earth would appear not to be at rest in space, but to be carried along by the earth. The theory of their experiment is as follows.

Let a pencil of light SA (Fig. 230), falling upon a piece of plane glass A, be partly reflected along AB and partly transmitted along AC. If the reflected and transmitted portions fall perpendicularly upon mirrors B and C they will be returned along BA and CA. Hence, if AB = AC,

[1] Michelson and Morley, *American Journal of Science* (3), vol. xxxi. p. 377, 1886.

[2] Michelson and Morley, *Phil. Mag.* vol. xxiv. p. 449, 1887.

the transmitted part of BA and the reflected part of CA will interfere along AD. Let us suppose now that the ether is at rest, and that the earth moves in the direction AC, so that the mirror A is carried to A′, while the reflected light travels to the mirror B and back again. In this case the ray SA will be reflected along AB′ where the angle BAB′ is equal to the aberration $a$. The reflected ray returns along B′A′, and the angle AB′A′ is equal to $2a$. The transmitted ray AC returns along CA′, and is reflected at A′ along A′D′, making CA′D′ $= 90 - a$, and therefore still coinciding in direction with the transmitted portion of B′A′. The returning rays B′A′ and CA′ do not, however, now meet at exactly the same point A′, but this difference is of the second order and negligible.[1]

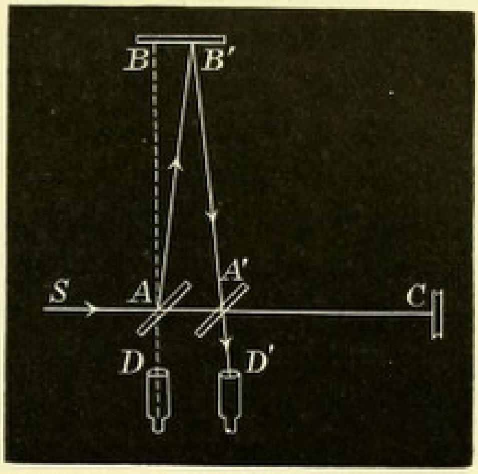



Fig. 230.

Let $v$ be the velocity of light, $u$ the velocity of the earth, D the distance AB or AC, T the time occupied by the ray in passing from A to C, and T′ the time in returning from C to A′. Then, since C is moving with velocity $u$, we have $v\mathrm{T} = \mathrm{D} + u\mathrm{T}$, so that

$$\mathrm{T} = \frac{\mathrm{D}}{v-u}, \quad \mathrm{T}' = \frac{\mathrm{D}}{v+u},$$

and the whole time is therefore

$$\mathrm{T} + \mathrm{T}' = 2\mathrm{D}\frac{v}{v^2 - u^2}.$$

The distance traversed in this time is therefore

$$2\mathrm{D}\frac{v^2}{v^2 - u^2} = 2\mathrm{D}\left(1 + \frac{u^2}{v^2}\right),$$

neglecting terms of the fourth and higher orders. The length of the other path AB′A′ is obviously

$$2\mathrm{D}\left(1 + \frac{u^2}{v^2}\right)^{\frac{1}{2}} = 2\mathrm{D}\left(1 + \frac{u^2}{2v^2}\right) \quad \text{(approx.)}$$

since AA′/AB $= 2u/v$.

The difference of the paths ACA′ and AB′A′ is consequently $\mathrm{D}u^2/v^2$. If the whole apparatus be now rotated through 90° the difference of path will be in the opposite direction, and a displacement of the interference fringes corresponding to the retardation

$$\delta = 2\mathrm{D}u^2/v^2.$$

[1] [For a more detailed account of the theory, see Larmor, *Aether and Matter*, pp. 46-53, Cambridge, 1900.]

should occur. Taking $u$ to be merely the velocity of the earth in its orbit, we have $u^2/v^2 = 10^{-8}$, and measuring D in wave lengths of yellow light Michelson and Morley found it was equal to $2 \times 10^7$ (about 11 metres) in their second experiment. Hence, if the ether be at rest—that is, if the relative motion of the earth with respect to it be $u$—we should have a displacement of the fringes equal to

$$4 \times 10^7 \times 10^{-8} = 0 \cdot 4 \text{ (of a fringe width).}$$

The actual displacement observed was certainly less than the twentieth part of this amount and probably less than the fortieth part.

Since the displacement $\delta$ is proportional to the square of the relative velocity $u$, the authors of the experiment conclude that the relative velocity of the earth and the ether is probably less than one-sixth, and certainly less than one-fourth of the earth's orbital velocity. They consequently regard it as tolerably certain that if there is any relative motion between the earth and the ether, it must be small,—so small, in fact, that the ordinary theory of aberration becomes untenable.

In conducting the experiment the chief difficulties encountered were the distortion of the apparatus produced by rotating it, and its extreme sensitiveness to vibration. The latter was so great that the interference fringes could not be observed when working in the city, except at brief intervals, even at two o'clock in the morning. These difficulties were surmounted by placing the apparatus on a massive stone floating on mercury, placed in a cast-iron trough which was cemented into a low brick pier. The stone rested on an annular wooden float, such that there was a clearance of about 1 centimetre between it and the sides of the iron trough. A pin kept the float concentric with the trough, and during the observations the apparatus was kept in slow uniform motion, making one revolution in about six minutes. This motion was slow enough to permit readily of the necessary observations, and the strains caused by bringing the system to rest at each observation were thus avoided.

The negative result obtained in the foregoing experiment, although being in accordance with the supposition that the ether in the neighbourhood of the earth is at rest relatively to the earth, or that the ether outside a moving body is dragged along viscously or otherwise by that body, may, nevertheless, be explained in some other manner. For example, it has been suggested by Professor FitzGerald that the force of attraction between two molecules moving through the ether in a given direction may depend on the angle which the line joining the molecules makes with the direction of motion.[1] If this be so, the force

[1] This idea has been also put forward by Professor H. A. Lorentz, and developed

will not be the same when the molecules are moving in the direction of the line joining them, as when they are moving in the perpendicular direction. By this means it is possible that a body which is spherical when at rest may become slightly ellipsoidal when moving; or, in other words, that the length of a given rod may depend upon the angle which the direction of its length makes with the direction of its motion through space. This change in the linear dimensions of Michelson and Morley's apparatus when rotated through a right angle may be such as to compensate the displacement of the fringes expected from the motion of the earth through the ether.

**316. Lodge's Experiment.**—The influence of moving matter on the velocity of light in its neighbourhood has been examined by Professor O. J. Lodge[1] in a recent series of experiments. The apparatus consisted of a pair of circular steel plates 3 feet in diameter. These plates were mounted with their planes parallel and 1 inch apart on a vertical axis, and were then set spinning at as high a speed as they would safely stand without flying to pieces. If the ether in the space between the spinning discs is dragged around with them, then a pencil of light travelling around between the discs in the direction of motion would be expected to traverse the space with a greater velocity than a beam travelling in the opposite direction. To test this a parallel beam of light was divided into two parts by a semi-transparent mirror (a piece of glass silvered so thinly that it transmits one-half of light and reflects the other half), and the two halves of this split beam were reflected by fixed mirrors in such a way that they passed in opposite directions round and round the space between the spinning discs, as shown in Fig. 231. The plane of this

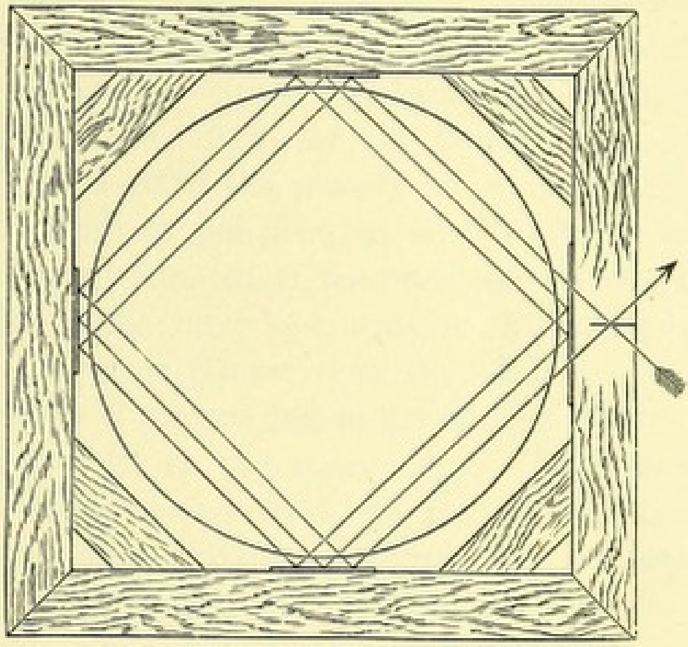


Fig. 231.

in a valuable memoir (*Versuch einer Theorie der electrischen und optischen Erscheinungen in bewegten Körpern*, Leiden, E. J. Brill, 1895).

[1] O. J. Lodge, "Aberration Problems," *Phil. Trans.* vol. clxxxiv. p. 727, 1893. In this paper several problems connected with aberration and the relative motion of matter and ether are discussed in detail.

figure is horizontal, and shows the space between the discs around which the beams were reflected several times by four fixed mirrors forming a square. The beams could thus be caused to traverse a distance of 30 or 40 feet between the discs, and then be allowed to enter a telescope, where they interfered and produced fringes.

At first a displacement of the fringes was obtained, but this proved to be spurious, being caused by the disturbing effect of the air-blast thrown off by the spinning discs. When due precautions were taken to avoid this spurious displacement it was concluded that there was no indication of a shift due to a viscous or other drag of the ether. Hence the velocity of light in the space just outside moving matter would appear to be uninfluenced by the motion of the matter—at least in the case of a small mass such as that used in these experiments.

## CHAPTER XX

### THE RAINBOW

**317. Introduction of the Geometrical Theory.**—Rainbows are seen when the sun shines upon falling rain or on the spray of a cascade or fountain. Sometimes only one bow is observed, but often two are seen, which are arcs of concentric circles, and have their common centre on the line joining the sun to the eye of the observer. These bows are seen only when the observer's back is turned towards the sun, and they have therefore from the earliest times been attributed to the refraction and reflection of the sun's light by the falling drops of water.[1] The inner bow, which is called the *primary bow*, is about 41° in angular radius. It is very brilliant and presents all the colours of the solar spectrum, the red being situated at its outer and the violet at its inner edge. The radius of the outer bow is about 52°, and it is termed the *secondary bow*. It is much fainter than the primary bow, and presents the same succession of colours, but in the reverse order—being red at the inner and violet at the outer edge. The space between the two bows is notably darker than the rest of the sky, which is pervaded by a general faint illumination inside the primary and outside the secondary bow.

[1] "This was understood by some of the antients, and of late more fully discussed and explained by the famous *Antonius de Dominis*, Archbishop of *Spalato*, in his book *De Radiis Visûs et Lucis*, published by his friend Bartolus at Venice in the year 1611, and written above twenty years before. For he teaches there how the interior bow is made in round drops of rain by two refractions of the sun's light, and one reflexion between them, and the exterior by two refractions and two sorts of reflexions between them in each drop of water, and proves his explications by experiments made with a phial full of water and with globes of glass filled with water, and placed in the sun to make the colours of the two bows appear in them. The same explication *Des-Cartes* hath pursued in his *Meteors*, and mended that of the exterior bow. But while they understood not the true origin of colours, it is necessary to pursue it here a little farther" (Newton, *Opticks*, book i. part ii. prop. ix.).

One of the "antients" referred to in the above extract was Theodorich, who, about 1311, anticipated the work of A. de Dominis. His writings were first published in 1814 by Venturi (see Tait's *Light*, p. 125).

After A. de Dominis the geometrical theory of the bows was taken up and developed by Descartes, but it was not until Newton had discovered the difference in refrangibility of the several colours, and their consequent separation by refraction, that the varied colour of the bow could be explained. The explanation of the colours as it left the hands of Newton, and the geometrical theory of the phenomena, afford only a first approximation to the solution of the problem. The primary and secondary bows, or, as we may call them, the principal bows, are accompanied by other coloured bands, or spurious bows, close to the inner edge of the primary and the outer edge of the secondary. These extra bands are called *supernumerary* or *complementary bows*, and their explanation is afforded only by the more rigorous application of the wave theory.

The complete theory of the rainbow is therefore not to be obtained from the simple consideration of rays, but must be treated as a phenomenon to be explained by the general principles of interference. This was first pointed out by Young, and the question was afterwards considered by Potter,[1] the solution being completed by Airy.[2] It will be convenient, however, to consider the principles of the geometrical theory before proceeding to the more complete solution.

**318. Deviation of a Ray emerging from a Refracting Sphere after any Number of Internal Reflections.**—The geometrical theory of the rainbow requires a knowledge of the manner in which a pencil of parallel light is refracted through a transparent sphere, and of the form of the emergent beam. Let SA (Fig. 232) be a ray of light falling on a refracting sphere ABC.

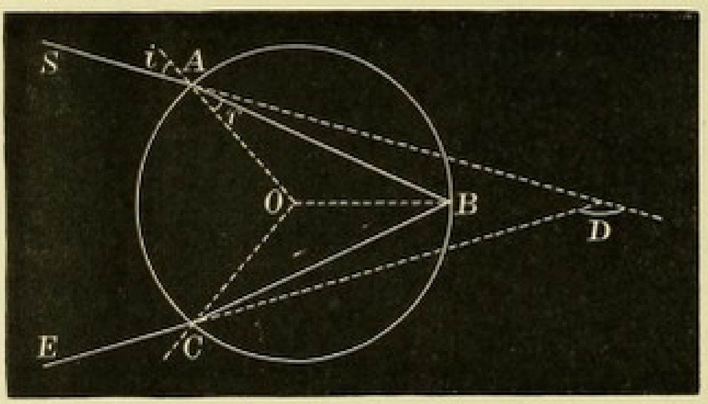


Fig. 232.

Let $i$ be the angle of incidence at A, and $r$ the angle of refraction, then the deviation produced at A is $i - r$. Again, since OA = OB = OC, it follows that OBA = OBC = OCB = $r$, consequently the deviation produced by reflection at B is $\pi - 2r$, and the deviation produced by refraction at C is as before $i - r$. The complete deviation of the ray CE emerging after a single internal reflection is therefore

$$D = 2(i - r) + \pi - 2r,$$

---

[1] Potter, *Trans. Camb. Phil. Soc.* vol. vi. p. 141.

[2] Airy, *Trans. Camb. Phil. Soc.* vol. vi. p. 379.

and it is obvious that, as each reflection introduces a deviation $(\pi - 2r)$, the deviation of a ray emerging after $n$ internal reflections will be

$$D = 2(i - r) + n(\pi - 2r) \qquad (1).$$

**319. Minimum Deviation.** — The foregoing general expression for the deviation shows that it will vary with the angle of incidence. We shall now prove that there is a certain incidence for which the deviation introduced is least, and the rays which suffer this least deviation are those with which we are most concerned in the theory of the bows. The change of D corresponding to a small variation of $i$ is by equation (1)

$$dD = 2di - 2(n+1)dr,$$

and for a maximum or minimum the value of D is stationary, that is $dD = 0$, which gives

$$di = (n+1)dr \qquad (2).$$

But since $\sin i = \mu \sin r$, we have also

$$\cos i\,di = \mu \cos r\,dr,$$

and therefore by (2)

$$\mu \cos r = (n+1) \cos i \qquad (3).$$

From this we obtain at once

$$\cos i = \sqrt{\frac{\mu^2 - 1}{n^2 + 2n}},$$

which determines the angle of incidence at which the deviation is stationary. It may be easily inferred that this incidence corresponds to the minimum deviation rather than to the maximum. For in the case of the ray passing through the centre, the deviation is $n\pi$, which is obviously greater than that of any other ray. However, by differentiating a second time we find

$$\frac{d^2D}{di^2} = -2(n+1)\frac{d^2r}{di^2}.$$

But

$$\frac{dr}{di} = \frac{\cos i}{\mu \cos r}, \quad \text{and} \quad \frac{d^2r}{di^2} = \frac{(1-\mu^2)\sin i}{\mu^3 \cos^3 r}.$$

This latter is always negative $(\mu > 1)$, and consequently the second derived of D is always positive—that is, the corresponding value of D is a minimum.

Substituting for $\mu$ in the preceding formula and taking $n$ equal to 1, 2, 3, 4 in succession, the corresponding minimum deviations are

| $n$ | red | violet |
|---|---|---|
| 1 | $\pi - 42°\cdot1$ | $\pi - 40°\cdot22$ |
| 2 | $2\pi - 129°\cdot2$ | $2\pi - 125°\cdot48$ |
| 3 | $3\pi - 231°\cdot4$ | $3\pi - 227°\cdot08$ |
| 4 | $4\pi - 317°\cdot07$ | $4\pi - 310°\cdot07$ |

**320. Dispersion.**—Since the refractive index is greater for the violet rays than for the red, the deviation of the violet after any number of refractions will be greater than that of the red, and we naturally infer that the least deviation of the red light emerging from a raindrop will be less than that of the violet. The same conclusion follows from the equation

$$D = n\pi + 2i - 2(n+1)r,$$

for $i$ is the same for all the colours and $r$ decreases from the red to the violet. The relation connecting a variation of the minimum deviation with the refractive index is easily determined, for we have

$$dD = 2di - 2(n+1)dr,$$
$$\cos i di = \mu \cos r dr + \sin r d\mu,$$
$$(n+1)\sin i di = \mu \sin r dr - \cos r d\mu \qquad \text{(by equation 3).}$$

Hence using the relation $\mu \cos r = (n+1)\cos i$, we have

$$\frac{dD}{d\mu} = \frac{2\sin r}{\cos i} = \frac{2}{\mu}\tan i = \frac{2}{\mu}\sqrt{\frac{(n+1)^2 - \mu^2}{\mu^2 - 1}}.$$

This expression being positive ($i < 90°$) shows that the minimum deviation increases with the refractive index.

**321. The Emergent Rays.**—Let ABC (Fig. 233) be a refracting sphere and consider a beam of parallel rays falling upon it in the

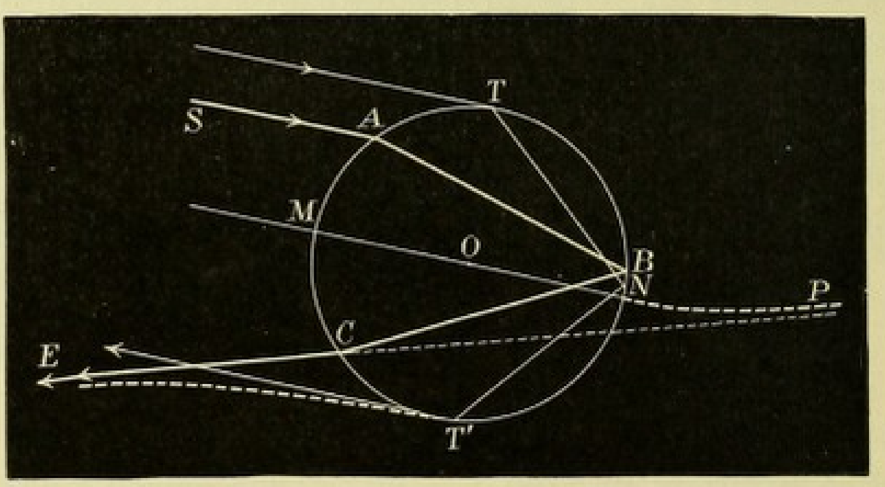



Fig. 233.

direction SA. Let SABCE be that ray which suffers the least deviation by the two refractions and a single internal reflection. The deviation being stationary at the minimum, it follows that in the direction CE there will emerge a pencil of approximately parallel rays, and in this direction there will be great concentration of illumination if there is no mutual interference of the rays. Any ray incident between M and A emerges, after reflection, from the arc MC, and its deviation is greater than that of EC; so also the rays which enter the sphere between A and T (the tangent ray) leave it by the arc T′C, and the emerging rays cut the ray CE, since they suffer a greater deviation than CE.

Thus all the rays emerging from the arc MT′ lie between the lines OM and CE. Hence the light which emerges after suffering a single reflection all lies within a cone of about 42°, and the bounding surface of the cone contains the rays which have suffered the least deviation. Outside this cone the eye will receive no light. In the direction CE, the edge of the cone, the illumination is very intense, and inside the cone the rays are divergent and the illumination feeble.

The rays emerging from the arc CT′ touch a caustic curve to which the ray CE is an asymptote and which touches the sphere at T′.

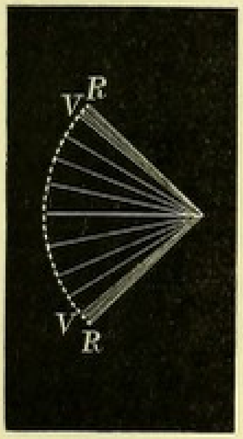



Fig. 234.—1 Reflection.

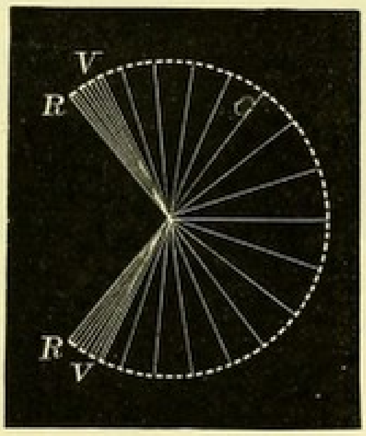



Fig. 235.—2 Reflections.

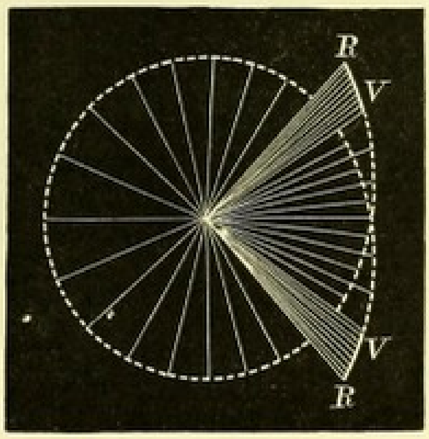



Fig. 236.—3 Reflections.

The rays emerging from the arc MC touch another branch NP, which is also asymptotic to EC produced backwards, and this branch meets the sphere orthogonally at N—that is, it touches the central ray MN. These properties of the caustic are easily deduced (Art. 324), but we shall first consider the formation of the bows. Preliminary to this consideration an inspection of diagrams (234, 235, 236), showing the minimum deviation rays and the emergent pencil for 1, 2, 3, internal reflections, will be useful.

**322. Formation of the Bows.**—The general explanation of the various bows is easily deduced from the foregoing principles. Let O, $O_1$, $O_2$, etc. (Fig. 237) be a series of raindrops in the same vertical line, or successive positions of the same drop, and let the observer's eye be situated at E. Draw EC parallel to the direction of the incident light. The line EC produced backwards will pass through the sun and will be the axis of the bows. Each drop emits light which has suffered one, two, three, etc., internal reflections, or, as we may call it, light of the first, second, and third, etc., orders. Let us first consider the emergent light of the first order. This emerges from each drop in a cone, of which the semi-vertical angle is about 41° for the violet and 43° for the red.

Consider the violet first and draw $EO_1$, making the angle $O_1EC$ equal to the supplement of the least deviation of the violet—that is, about 41°. Then the drop $O_1$ will send its least deviated rays to E, and will consequently appear very luminous. Any drop, such as O situated below $O_1$, will also send light to E, but in this case it will be from the interior of the cone, the rays reaching E from O will be divergent, and the illumination very faint. On the other hand, drops above $O_1$ will fail to transmit any light of the first order to E, for the point E is situated entirely outside the cones emerging from them.

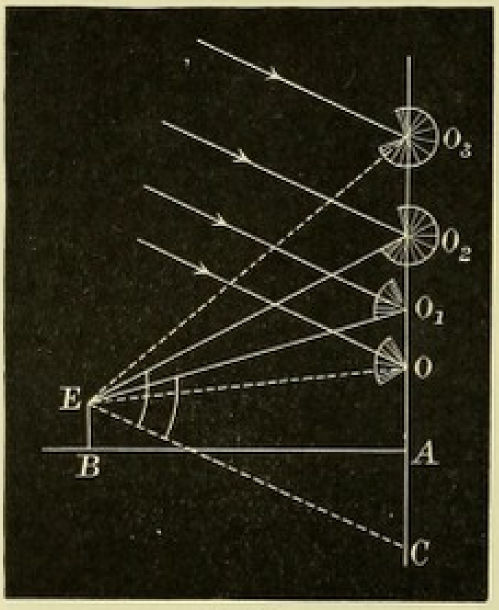


Fig. 237.

If now the line $EO_1$ be made to revolve round EC, keeping the angle between them constant, it will generate a right cone of which EC is the axis, and the point $O_1$ will describe a circle. Every drop on this circle will obviously transmit minimum deviation light of the first order to E, drops inside the circle will transmit faint illumination, and drops outside it will send no light to the eye.

What we have, consequently, is a brilliant circle, the space inside of which is faintly illuminated and the outside dark. On account of the sensible diameter of the sun the bright circle will not be merely a bright line, but will have a sensible width determined by the apparent diameter of the sun. The least deviation of the red being about 43°, it follows that the drops which transmit the red will be a little higher up than those which give the violet. The angular radius of the red circle will be 43°, and the other colours will occupy intermediate positions. We have therefore a spectrum-coloured band about 2° in width, red at the outer and violet at the inner edge.

Since each simple colour gives rise to a band of definite width, on account of the magnitude of the sun's disc, it follows that overlapping and considerable mixture of colours is presented in the rainbow. This may take place to such an extent that all trace of colour may become obliterated, or nearly so. We have then the phenomenon of the *white rainbow.* This may happen when the sun shines on the raindrops through a thin cloud in the higher regions of the atmosphere. The sky is then very bright for one or two degrees around the sun, and the diameter of the effective source of light is very much increased.

Let us now consider the light which suffers two internal reflections. The inclination to the incident light of the least deviated red ray is about 51° and that of the violet about 54° (see Fig. 235). Hence, if the angle $CEO_2$ is equal to 51°, the drop $O_2$ will send intense red light of the second order to the eye, and drops a little above $O_2$ will transmit the other colours. We will thus have a second bow, produced by the light of the second order. This bow will be broader and fainter than the primary, for the light is weakened by the second interior reflection, and the dispersion is also greater, on account of the lengthened path within the drop. Drops below $O_2$ transmit no second order light to E, and those above $O_2$ illuminate E faintly. It consequently follows that the space between the two bows is dark compared with the spaces inside the primary and outside the secondary. The width of the secondary bow is about 3°, its outer edge is violet and its inner red.

In the same manner we have bows arising from three, four, or more internal reflections. The third and fourth bows are situated between the observer and the sun, so that to be seen the rain should be viewed with the eye directed towards the sun (see Fig. 236). The brightness of the sun's light is, however, sufficient to render their faint illumination unobservable. The fifth bow, however, occurs in the same part of the sky as the first and second, and would be seen by an observer with his back turned towards the sun but for its extreme faintness. Its light is so much weakened by the five reflections, and dispersion, that it is rarely, if ever, detected.

Intersecting rainbows are also sometimes observed. They require for their production two sources of parallel rays, such as the sun and its image by reflection in a large sheet of calm water behind the observer. The usual system of bows is formed by the direct light of the sun, and the second system, intersecting them, by the reflected light, which proceeds, as it were, from the image of the sun beneath the water. The second system is like the ordinary system, but is less brilliant, and the bows have their common centre as much above the horizon as that of the direct system is below it.

Lunar bows may also be seen under favourable circumstances. They are, however, faint and their colours are not easily distinguished, except the moon be full and other conditions favourable. This, together with the fact that they occur only at night, renders their observation rare.

Much foolish discussion seems to have been raised by the question whether two persons, or even the two eyes of the same person, see the same rainbow. From what has been said it is plain that the system of raindrops which transmits rays of minimum deviation to one point cannot be the same as that which sends them to another. The ring of drops sending light of least deviation to E lies on a cone of which E is the vertex, EC the axis, and the semi-vertical angle equal to the supplement of the least deviation. The question might as well be asked as to whether two persons see the same object by means of the same rays of light.

A similar question is, Can a rainbow be seen by reflection?—that is, whether a bow apparently reflected in still water is the *image* of that seen directly in the sky. In this case the reflected bow is not the image of that seen in the sky, but is the reflection of that which would be seen by an eye vertically below that of the observer, and as much below the surface of the water as his eye is above it.

**323. Supernumerary Bows.**—The geometrical theory established by Descartes and Newton was for a long time accepted as affording a complete and satisfactory explanation of the phenomenon of the rainbow. Closer observation, however, showed that in addition to the principal bows already described, there are other concomitant coloured bands, or alternations of brightness and darkness, and of these the elementary theory gives no account. These additional bands are frequently seen near the inner edge of the primary bow and less frequently at the outer edge of the secondary. They are called *supernumerary*, *complementary*, or *spurious bows*, and are best defined near their summits.

Young[1] was the first to propose an explanation of these spurious bows on the theory of interference, and Mr. Potter[2] afterwards proved that the complete theory of the rainbow is to be obtained not from the geometrical theory of rays, but from the principle of interference. Young remarked that the efficacious rays consisted of two superposed sets, parallel in direction on emergence, but traversing different paths in the drop. Thus the rays emerging from the drop near the ray CE (Fig. 233) consist of two sets, viz. those which escape from the lower arc CT′ and those which escape from the upper arc CM. The former will obviously be parallel to some of the latter, and as they travel over different paths in the drop destructive interference is to be expected when they are in opposite phases. Near the direction of the minimum deviation EC we should expect not merely a single maximum

[1] Young, *Phil. Trans.* p. 8, 1804.

[2] Potter, *Camb. Phil. Trans.* vol. vi. p. 141.

of brightness, but a maximum accompanied by alternations of bright and dark bands.

To complete the solution we require the form of the wave emerging from the drop. This was effected by Mr. Potter,[1] who first traced the caustic enveloped by the emergent rays and then developed the wave front as its involute. The final investigation and the determination of the intensity at any point was given by Airy.[2]

**324. Caustic of the Emergent Rays.**—Let CE (Fig. 238) be one of the emergent rays, D its deviation, and P the point where it intersects the consecutive ray. Then P is a point on the caustic. Denote OP by $\rho$, the radius of the circle by $a$, and draw OQ perpendicular to CP. We have, if OPQ $= \gamma$,

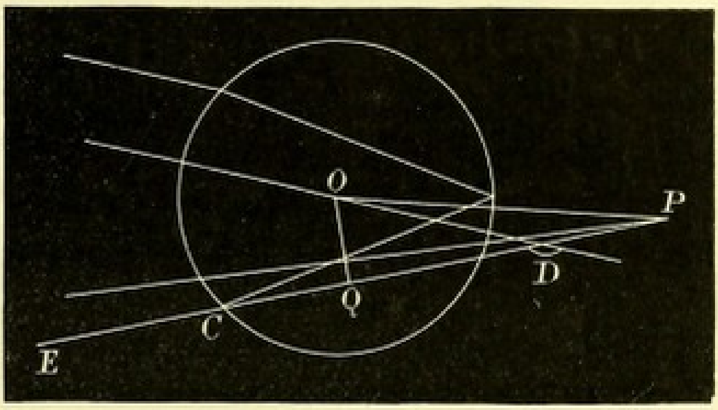



Fig. 238.

$$OQ = a \sin i = \rho \sin \gamma \quad (1),$$

and since P is the intersection of two consecutive rays, $\rho$ remains constant, while $i$ and $\gamma$ change to $i + di$ and $\gamma + d\gamma$, therefore

$$a \cos i di = \rho \cos \gamma d\gamma,$$

or by means of equation (1)

$$\frac{d\gamma}{di} = \frac{a \cos i}{\rho \cos \gamma} = \frac{\tan \gamma}{\tan i} \qquad (2).$$

But

$$D = 2(i - r) + n(\pi - 2r).$$

Therefore

$$\frac{dD}{di} = 2 - 2(n+1)\frac{dr}{di} = 2\left\{1 - (n+1)\frac{\cos i}{\mu \cos r}\right\} \qquad (3),$$

since $\sin i = \mu \sin r$.

Now it is clear that $\frac{dD}{di} = -\frac{d\gamma}{di}$, therefore by (2) and (3)

$$\tan \gamma = -2 \tan i \left\{1 - (n+1)\frac{\cos i}{\mu \cos r}\right\} \qquad (4),$$

and therefore by (1) and (4)

$$\rho = \frac{a \sin i}{\sin \gamma} = a \sin i \sqrt{1 + \cot^2 \gamma}$$

$$= a \sqrt{\sin^2 i + \frac{\mu^2 \cos^2 i \cos^2 r}{4\{\mu \cos r - (n+1) \cos i\}^2}}.$$

---

[1] Potter, *loc. cit.*

[2] Airy, *Camb. Phil. Trans.* vol. vi. p. 379.

The value of $\rho$ becomes infinite when

$$\mu \cos r - (n+1)\cos i = 0.$$

That is, when

$$\cos i = \sqrt{\frac{\mu^2 - 1}{n^2 + 2n}}.$$

The direction of the least deviated rays is therefore asymptotic to the caustic. Again, if $i = 90^\circ$, we have $\rho = a$, so that the caustic meets the drop at the point T′, where the light emerges tangentially (Fig. 233).

In the case of a single internal reflection we have $n = 1$, and if we take $\mu = \frac{4}{3}$, as is the case for water, we find that $\rho$ is also equal to $a$ when $i = 0$—that is, the caustic meets the drop at the point N where the light falls normally on the sphere.

The curve for a single internal reflection is shown in Fig. 233. The rays emerging from the arc CT′ touch the branch ET′, and those emerging from CM touch the branch NP, while the ray of least deviation CE meets both branches at infinity.

The form of the wave front may be derived from the caustic, for the rays are normals to the wave front and tangents to the caustic. The wave front is therefore an involute of the caustic. For the branch NP this will be a curve OR (Fig. 239), and for the branch ET′ a curve OR′. The asymptotic ray CE will be normal to both OR and OR′ at O, and the point O will be a point of inflection, the tangent at which will be perpendicular to the ray CE.

**325. Intensity at any Point—Airy's Investigation.**—In order to determine the intensity at any point we require the equation of the surface of the emergent wave. Let ROR′ (Fig. 239) be a section of the wave front by the plane of incidence, OX the least deviated ray, and OY the inflectional tangent to the curve at O. Taking OX and OY for axes of reference, we have $\frac{dx}{dy} = 0$, and $\frac{d^2x}{dy^2} = 0$ at the origin. Consequently near O the form of the curve will be sufficiently represented by the equation

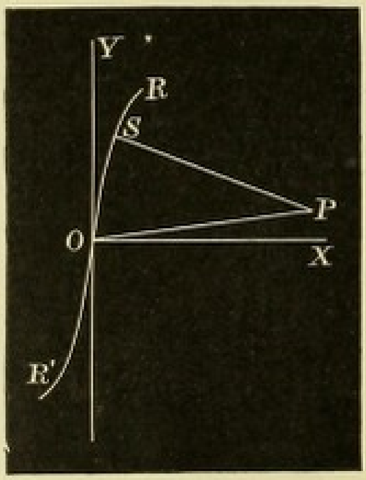


Fig. 239.

$$x = Ay^3,$$

and in calculating the intensity at any point near the axis OX we need only consider a portion of the wave in the neighbourhood of the origin.

Let P be any point near the axis OX, so that its ordinate $y$ is small compared with its abscissa $x$. Let S be a point on the wave front, $s$ the arc OS. Then if the vibration propagated to P by the element at O be $ds \,.\, \sin \omega t$, that transmitted by the element at S will be $ds \,.\, \sin(\omega t + \delta)$ where $\delta$ is the phase difference, and the complete disturbance at P will be

$$\int_{-\infty}^{+\infty} \sin(\omega t + \delta) dy,$$

for near O the curve closely approximates to the axis OY, and $ds$ is equal to $dy$. It remains to calculate $\delta$.

Let the co-ordinates of S be $x$ and $y$, those of P, $\xi$ and $\eta$. Then if $PS = r$ we have

$$r^2 = (x-\xi)^2 + (y-\eta)^2 = y^2 - 2y\eta + \eta^2 + A^2y^6 - 2Ay^3\xi + \xi^2,$$

substituting $Ay^3$ for $x$ from the equation of the curve. Replacing $\xi^2 + \eta^2$ by $\rho^2$ and neglecting $y^6$, since $y$ is small, we have

$$\begin{aligned} r &= (\rho^2 - 2\eta y + y^2 - 2A\xi y^3)^{\frac{1}{2}}, \\ &= \rho\left(1 - \frac{\eta y}{\rho^2} + \frac{y^2}{2\rho^2} - \frac{A\xi y^3}{\rho^2} - \frac{\eta^2 y^2}{2\rho^4} + \frac{\eta y^3}{2\rho^4} - \frac{\eta^3 y^3}{2\rho^6}\right), \\ &= \rho - \frac{\eta y}{\rho} + \frac{\rho^2 - \eta^2}{2\rho^3} y^2 + \left\{\frac{\eta(\rho^2 - \eta^2)}{2\rho^5} - \frac{A\xi}{\rho}\right\} y^3. \end{aligned}$$

Now $\eta$ is small compared with $\xi$, so that $\rho = \sqrt{\xi^2 + \eta^2} = \xi + \frac{\eta^2}{2\xi}$, and the expression for $r$ becomes

$$\begin{aligned} r &= \xi + \frac{\eta^2}{2\xi} - \frac{\eta y}{\xi} + \frac{y^2}{2\xi} + \left(\frac{\eta}{2\xi^3} - A\right) y^3, \\ &= \xi + \frac{\eta^2}{2\xi} - \frac{\eta y}{\xi} + \frac{y^2}{2\xi} - Ay^3 \end{aligned}$$

approximately.

The term involving the second power of the variable may be removed by writing

$$y = z + \frac{1}{6A\xi},$$

and we then have, writing $c$ for the absolute term,

$$\begin{aligned} r &= c - \frac{(4\xi\eta - 1/3A)}{4\xi^2} z - Az^3, \\ &= c - \frac{\eta}{\xi} z - Az^3, \end{aligned}$$

neglecting $1/3A$, which is of the order $y^3/x$, in comparison with $4\xi\eta$. The expression for the resultant disturbance is therefore

$$\int_{-\infty}^{+\infty} \sin\left\{\omega t + \frac{2\pi c}{\lambda} - \frac{2\pi}{\lambda}\left(\frac{\eta z}{\xi} + Az^3\right)\right\} dz,$$

and the intensity is proportional to

$$\left[\int_{-\infty}^{+\infty} \cos \frac{2\pi}{\lambda}\left(\frac{\eta z}{\xi} + Az^3\right) dz\right]^2 + \left[\int_{-\infty}^{+\infty} \sin \frac{2\pi}{\lambda}\left(\frac{\eta z}{\xi} + Az^3\right) dz\right]^2.$$

The second integral is zero, and the first is twice as great as if the limits were zero and infinity, consequently

$$I = 4\left[\int_0^\infty \cos\frac{2\pi}{\lambda}\left(\frac{\eta z}{\xi} + Az^3\right)dz\right]^2.$$

This integral may be written in the form

$$\int_0^\infty \cos\frac{\pi}{2}(w^3 + mw)dw,$$

where

$$w = z\left(\frac{4A}{\lambda}\right)^{\frac{1}{3}}, \quad \text{and } m = \frac{4}{\lambda}\frac{\eta}{\xi}\left(\frac{\lambda}{4A}\right)^{\frac{1}{3}}.$$

The values of this integral have been calculated by Airy for successive values of $m$ by a method of approximation analogous to that adopted by Fresnel in the evaluation of his integrals.

As $m$ increases, the integral attains a maximum, and then passes in succession through a series of minima and maxima, the first maximum being much greater than any of its successors. The zero value of $m$, which corresponds to $\eta = 0$—that is, for points on the line OX—does not then give a maximum, but the first maximum occurs at a point where $\eta/\xi$ has a positive value greater than zero.

It follows therefore that the first bright band—that is, the rainbow—is not situated exactly on the line of least deviation, but is attained at a deviation a little greater, or the angular radius of the primary bow is a little less than that indicated by the geometrical theory, and inside the bow we have several alternations of brightness and darkness—that is, supernumerary or spurious bands.

The angular distances between the spurious bands and the principal bow vary with the diameters of the raindrops. For the integral which determines the intensity reaches a maximum or a minimum for definite values of $m$, consequently the corresponding ratio $\eta/\xi$ which determines the position of the band varies directly as $(A)^{\frac{1}{3}}$. But two drops of different diameters, $a$ and $a'$, give emergent waves of which the sections are similar, and if $x, y; x', y'$, be homologous points on them we should have

$$\frac{a}{a'} = \frac{x}{x'} = \frac{y}{y'}.$$

But

$$\frac{x}{x'} = \frac{Ay^3}{A'y'^3} \quad \therefore \frac{y^2}{y'^2} = \frac{A'}{A} = \frac{a^2}{a'^2},$$

and

$$\left(\frac{A}{A'}\right)^{\frac{1}{3}} = \left(\frac{a'}{a}\right)^{\frac{2}{3}},$$

hence $\eta/\xi$ varies inversely as $a^{\frac{2}{3}}$—that is, the smaller the drops the farther apart are the supernumerary bows; and this explains why they are only well defined near their summits, for as the drops descend they grow in magnitude, and consequently the lower portions of the bow deviate less and less from the geometrical position.

The same reasoning applied to the second principal bow shows that it also is accompanied by spurious bands situated at its outer edge and best defined at its summit. This bow also deviates from the geometrical position, its radius being somewhat greater than that assigned by the elementary theory.

**326. Miller's Experiments.**—Airy's theoretical results have been confirmed by the experimental investigations of Mr. Miller.[1] A pencil of sunlight was admitted in a horizontal direction through a narrow vertical slit, and fell upon a thin vertical jet of water. Viewing the stream through a telescope (or with the naked eye) portions of the primary and secondary bows, and a large number of the spurious bands could be seen forming a series of vertical coloured fringes arranged side by side. The diameter of the water jet was about ·022 of an inch. The mixture of colours rendered it difficult to fix upon the brightest parts of the bands. The mean of eight observations was as follows:—

| | |
|---|---|
| Radius of brightest part of primary bow | 41° 32′ |
| ,, ,, ,, first spurious bow | 40° 27′ |
| Radius of brightest part of secondary bow | 51° 58′ |
| ,, ,, ,, first spurious bow | 53° 57′ |

According to the elementary theory we should have

| | |
|---|---|
| Radius of brightest part of primary bow | 41° 53′·9 |
| ,, ,, ,, secondary bow | 51° 12′·9 |

In conclusion it may be remarked that the light of the rainbow is partially polarised. This polarisation was noticed by Biot as early as 1811, and is to be expected as a consequence of the reflection and refraction suffered by the light in the drops of rain (Art. 173). The extent to which the light of any bow is polarised may be easily calculated by resolving the incident vibration into two components—one polarised in the plane of incidence, and the other perpendicular to it, and then applying the formulæ of Arts. 208, 209. Light polarised.

[1] Miller, *Camb. Phil. Trans.* vol. vii. p. 277.

## CHAPTER XXI

### ELECTROMAGNETIC RADIATION

**327. Ether demanded by Electric Phenomena — An Electric Charge a Charge of Energy and an Electric Current a Flow of Energy.**—To account for the propagation of heat and light—that is, of radiant energy—we have postulated the existence of a medium filling all space. But the transference of the energy of radiant heat and light is not the only evidence we have in favour of the existence of an ether. Electric, magnetic, and electromagnetic phenomena (and gravitation itself) point in the same direction.

It is a matter of common observation that attractions and repulsions take place between electrified bodies, magnets, and circuits conveying electric currents. Large masses may be set in motion in this manner and acquire kinetic energy. If an electric current be started in any circuit, corresponding *induced* currents spring up in all neighbouring conductors; yet there is no visible connection between the circuit and the conductors. To orginate a current in any conductor requires the expenditure of energy. How then is the energy propagated from the circuit to the conductors? If we believe in the continuity of the propagation of energy—that is, if we believe that when it disappears at one place and reappears at another, it must have passed through the intervening space, and therefore have existed there somehow in the meantime — we are forced to postulate a vehicle for its conveyance from place to place, and this vehicle is the ether.

When a body is electrified, what we must first observe is that a certain amount of energy has been spent; work has been done, and the result is the electrified state of the body. The process of electrifying a conductor is therefore the storing of energy in some way in, or around, the conductor in some medium (the ether). The work is spent in altering the state of the medium, and when the body is discharged the ether returns to its original state, and the store of energy is evolved.

Similarly a supply of energy is required to maintain an electric current, and the phenomena arising from the current are manifestations of the presence of this energy in the ether around the circuit. Formerly an electrified body was supposed to have something called electricity residing upon it which caused the electrical phenomena, and an electric current was regarded as a flow of electricity travelling along the wire, while the energy which appeared at any part of the circuit (if considered at all) was supposed to have been conveyed along the wire by the current. The existence of induction, however, and electromagnetic actions between bodies situated at a distance from each other, lead us to look upon the medium around the conductors as playing a very important part in the development of the phenomena. It is, in fact, the storehouse of the energy.

Upon this basis Maxwell founded his theory of electricity and magnetism, and determined the distribution of the energy in the various parts of the field in terms of the electric and magnetic forces. The ether around an electrified body is charged with energy, and the electrical phenomena are manifestations of this energy, and not of an imaginary electric fluid distributed over the conductor. When we speak of the charge of an electrified conductor we refer to the charge of energy in the ether around it, and when we talk of the electric flow or current in a circuit we refer to the only flow we know of, viz. the flow of energy through the electric field into the wire.

**328. Polarisation of the Ether—Electric Waves.**—The work spent in producing the electrification of a conductor is spent on the ether and stored there, probably as energy of motion. To denote this we shall say that the ether around the conductor is *polarised*, this word being employed to denote that its state or some of its properties have been altered in some manner by the work done on it—that is, by the energy stored in it. In the case of a conductor possessing what is termed a positive charge, the ether around it is polarised in a certain manner and to a certain extent depending on the intensity of the charge. If the charge be negative the polarisation is in the opposite sense, the two being related, perhaps, like right-handed and left-handed twists or rotations.

Now consider the case of a body charged alternately, positively and negatively, in rapid succession. The positive charge means a positive polarisation of the ether, which begins at the conductor and travels out through space. When the body is discharged the ether is once more set free and resumes its former condition. The negative charge now entails a modification of the ether or polarisation in the opposite sense. The result of alternate charges of opposite sign is that the ether at any point becomes polarised alternately in opposite directions, while waves of opposite polarisations are propagated through space, each carrying energy derived from the source or agent supplying the electrification. Here, then, we have a periodic disturbance of some kind occurring at each point, accompanied by waves of energy travelling outwards from the conductor.

The phenomena of interference lead to the conclusion that light is the result of a periodic disturbance, or vibration, of the ether, but as to the nature of the vibration—that is, as to the exact nature of the periodic change—or what it is that changes, we possess no knowledge. From the foregoing we see that alternating electric charges are accompanied by corresponding changes of state, or vibrations, of the ether, and if the charge be varied periodically and with sufficient rapidity, we have a vibration at each point analogous to, and perhaps identical with, that which occurs in the propagation of light.

This, then, is the electromagnetic theory of the luminous vibration. In the older or elastic-solid theory, the light vibrations were supposed to be actual oscillations of the elements or molecules of the ether about their positions of rest, such as takes place when waves of transverse disturbance are propagated through an elastic solid. Such a limitation is, however, unwarranted. All we know is that the change, disturbance, vibration, polarisation, or whatever we wish to term it, is periodic and transverse to the direction of propagation. The electromagnetic theory teaches us nothing further as to its nature, but rather asserts that whatever the change may be, it is the same in kind as that which occurs in the ether when the charge of an electrified body is altered or reversed. It reduces light and heat waves to the same category as waves of electric polarisation; the only quality of the latter required to constitute the former is sufficient rapidity of alternation. These speculations have received the strongest confirmation by the recent important experiments of Professor Hertz. Before describing them we shall consider the mode of discharge of a condenser. The theoretical investigation was given by Sir William Thomson[1] (Lord Kelvin) as early as 1853.

**329. Oscillating Discharge.**—When any elastic substance is subjected to strain and then set free, one of two things may happen. The substance may slowly recover from the strain and gradually attain its natural state, or the elastic recoil may carry it past its position of equilibrium, and cause it to execute a series of oscillations. Something of the same sort may also occur when an electrified condenser such as a Leyden jar is discharged. In ordinary language there may be a

[1] Sir William Thomson, *Phil. Mag.* June 1853.

continuous flow of electricity in one direction till the discharge is completed, or an oscillating discharge may occur—that is, the first flow may be succeeded by a back-rush, as if the first discharge had overrun itself and something like recoil had set in. The jar thus becomes more or less charged again in the opposite sense, and a second discharge occurs, accompanied by a second back-rush, the oscillation going on till all the energy is either radiated or used up in heating the conductors.

Let Q be the charge of the jar at any instant, C its capacity, R the resistance of the circuit, and L its coefficient of self-induction. Then if I be the intensity of the current and E the electromotive force, we have the equation

$$E - IR = \frac{d}{dt}(LI) = L\frac{dI}{dt}.$$

In this case $E = Q/C$, and $I = -\frac{dQ}{dt}$. Therefore

$$L\frac{d^2Q}{dt^2} + R\frac{dQ}{dt} + \frac{Q}{C} = 0.$$

The solution of this equation is

$$Q = Ae^{\mu t} + Be^{\mu' t},$$

where $\mu$ and $\mu'$ are the roots of the equation

$$\mu^2 + \frac{R}{L}\mu + \frac{1}{CL} = 0,$$

or

$$\mu = -\frac{R}{2L} \pm \sqrt{\frac{R^2}{4L^2} - \frac{1}{CL}}.$$

Writing

$$\alpha = \sqrt{\frac{R^2}{4L^2} - \frac{1}{CL}},$$

we have

$$\mu = -R/2L + \alpha, \quad \mu' = -R/2L - \alpha$$

and

$$Q = e^{-\frac{Rt}{2L}}(Ae^{\alpha t} + Be^{-\alpha t}),$$

where A and B are constants determined by the initial conditions, viz. that initially we have $Q = Q_0$, and $I = 0$, which give

$$A + B = Q_0, \quad \text{and } A\mu + B\mu' = 0,$$

or

$$A = Q_0\left(\frac{1}{2} + \frac{R}{4L\alpha}\right), \quad \text{and } B = Q_0\left(\frac{1}{2} - \frac{R}{4L\alpha}\right).$$

Hence at any time we have

$$Q = Q_0 e^{-\frac{Rt}{2L}}\left\{\left(\frac{1}{2} + \frac{R}{4L\alpha}\right)e^{\alpha t} + \left(\frac{1}{2} - \frac{R}{4L\alpha}\right)e^{-\alpha t}\right\}.$$

Consequently the current at any instant is

$$I = -\frac{dQ}{dt} = \frac{Q_0}{2CLa} e^{-\frac{Rt}{2L}} \left( e^{at} - e^{-at} \right).$$

Hence if $a$ be real—that is, if we have $R^2 > 4L/C$—the quantity Q will gradually diminish to zero as the time increases.

If, however, we have $R^2 < 4L/C$, then $a$ will be imaginary, and writing

$$a' = a\sqrt{-1} = \sqrt{\frac{1}{CL} - \frac{R^2}{4L^2}},$$

the above formulæ become at once

$$Q = Q_0 e^{-\frac{Rt}{2L}} \left( \cos a't + \frac{R}{2La'} \sin a't \right),$$

and

$$I = \frac{Q_0}{CLa'} e^{-\frac{Rt}{2L}} \sin a't.$$

In this case the current starts from zero and rises to a maximum; it then falls to zero and becomes reversed, after which it passes through a series of oscillations. The discharge therefore does not take place in a single flow from one coating to the other, but a back-rush sets in, and a series of currents, or oscillations, occur alternately in opposite directions.

The current attains its maximum intensity when

$$\tan a't = 2La'/R \qquad \text{(maximum current)}.$$

The zero value of the current is reached when

$$a't = n\pi \qquad \text{(zero current)},$$

and consequently the charge at the same time is at its maximum, for we have $I = -dQ/dt$. Thus the charge oscillates backwards and forwards, attaining positive and negative maxima after the lapse of equal intervals $\pi/a'$, the time of a complete oscillation being

$$T = \frac{2\pi}{\sqrt{\frac{1}{CL} - \frac{R^2}{4L^2}}}.$$

If the resistance be small compared with the reciprocal of the capacity we may use the approximate formula[1]

$$T = 2\pi\sqrt{CL}.$$

[1] If the capacity be expressed in electrostatic measure, and the self-induction in electromagnetic, this expression takes the form $2\pi\sqrt{CL}/v$ where $v$ is the velocity of light.

The successive maximum charges occur when $I=0$, or $\alpha' t = n\pi$; they are therefore

$$Q_0, \quad Q_1 = -Q_0 e^{-\frac{\pi R}{2L\alpha'}}, \quad Q_2 = Q_0 e^{-\frac{2\pi R}{2L\alpha'}}, \quad Q_3 = -Q_0 e^{-\frac{3\pi R}{2L\alpha'}}.$$

The quantities therefore diminish in geometrical progression, and the energy of the charge diminishes correspondingly on each oscillation, being lost by radiation into space, or in heating the circuit, or both.

Whether the discharge is continuous or oscillatory therefore depends on whether $4L$ is less or greater than $CR^2$, and an oscillatory discharge may be obtained either by increasing L or sufficiently diminishing C and R.

These predictions of analysis have been confirmed, as Thomson suggested, by examining the spark, during discharge, by means of a revolving mirror. In Feddersen's experiments the image of the spark in a revolving mirror was viewed through a telescope. When the resistance of the circuit was high the spark was merely drawn out in width—that is, at right angles to its length; but when the resistance was sufficiently reduced, so that the oscillating discharge might occur, the band was reduced to a broken image consisting of a series of strips, each strip corresponding to a discharge.[1]

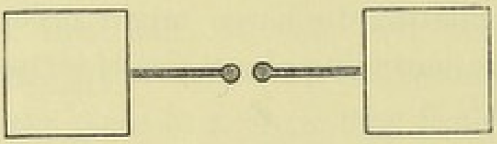
Paalzow examined the discharge through a vacuum tube in the same manner. With a high resistance a bluish light showed itself at one pole, but with a low resistance it was exhibited at both. In the latter case a magnet held near the tube split the discharge band into two lines of light, but in the former the magnet effected no separation of the discharge, showing that the discharge occurred in one direction only or in both, according as the resistance of the circuit was high or low.

The time of oscillation of an ordinary Leyden jar is from $\frac{1}{6}$ to $\frac{1}{3}$ millionth of a second, giving rise to electromagnetic waves about 50 to 100 metres long, if we assume their rate of propagation to be the same, or approximately the same, as that of light. A small thimble-sized jar would give rise to waves about 1 metre in length, and if we push the reduction to the limit of the light waves, or 6000 tenth-metres, we would require a circuit approaching atomic dimensions. This suggests that the long electromagnetic waves emitted by a condenser undergoing oscillating discharge are in reality the same in kind as the short ethereal waves which affect the retina, and that the light and heat waves are excited by electric oscillations in the atoms or molecules of the incandescent matter.

[1] For an account of the researches made in this department see Wiedemann's *Elektricität*, vol. iv. pp. 177, etc.

**330. Hertz's Experiments.**—The electromagnetic waves radiated by a condenser undergoing oscillating discharge have been ingeniously detected by Professor Hertz. The method is analogous in principle to the use of resonators in the detection of sound waves. It is well known that a vibrating tuning-fork when held near an open pipe will throw the air column into vibration and elicit a note from the pipe if the length of the pipe be so adjusted that its period is the same as that of the fork. In the same way, if an oscillating discharge be maintained in a circuit, it will play the part of the tuning-fork, and if another circuit be at hand which possesses the same periodic time of electric oscillation, then it will play the part of the resonator, and will be thrown into electric vibration. These sympathetic vibrations have been detected and examined by Hertz as follows.

The vibrator consisted of two brass plates A and B (Fig. 240), to each of which is attached a stiff wire terminated by a brass knob. The knobs are gilt and placed about 2 or 3 millimetres apart, so that when the plates are oppositely electrified a spark can easily cross between them. Electric oscillation then sets up, the charge passing backwards and forwards from one plate to the other and diminishing continually in quantity as its energy becomes used up in heating the wires, sparking, and especially by radiation into space. The time of each oscillation is approximately $2\pi\sqrt{CL}$, and this in the experiment will be about 1/30,000,000 of a second, if the plates used are about 40 centimetres square.



A Fig. 240.—The vibrator. B

At each passage of the charge between the plates of the vibrator, induced currents will occur in neighbouring conductors, and if the periodic time of an oscillation in one of these should happen to be the same as that of the vibrator, the oscillations induced in it will become multiplied and may attain a considerable intensity. This conductor then acts as a resonator. The resonator (or receiver) devised by Hertz was a simple circle of wire (Fig. 241), the ends terminating in brass knobs which could be adjusted at a small distance apart. The length of the wire[1] being carefully adjusted to suit the period of oscillation of the vibrator, then, when the vibrator is in action, the induced current flows backwards and forwards in the resonating circle from

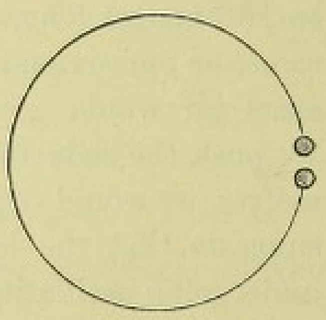


Fig. 241.—The receiver.

[1] If the plates of the vibrator are 40 cm. square, the length of the receiver is about 210 cm. of No. 17 wire.

one knob to the other, and finally attains such a strength that sparking actually occurs between the knobs.

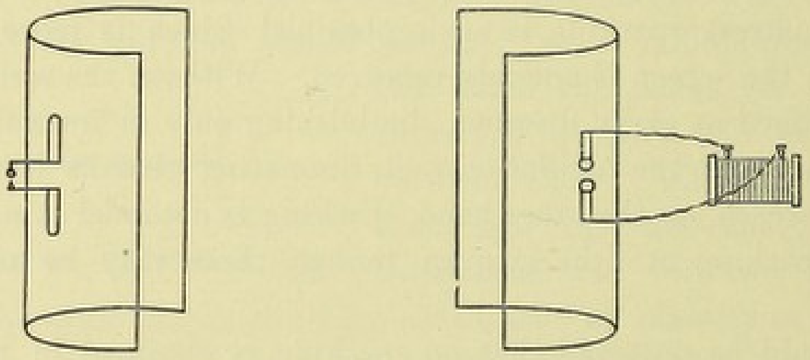
The charging of the vibrator is effected by connecting its plates with the terminals of an induction coil. In this way the oscillation is maintained in the vibrator, and continuous sparking can be obtained in the resonating circle.

The reflection and interference of electromagnetic waves may now be observed with facility. Placing a large metal screen (a sheet of zinc 2 or 3 metres square) immediately behind the resonator, the sparking is observed to increase in brilliancy, and the distance between the knobs may be augmented beyond the limiting distance at which sparking would occur before the screen was brought up. When the screen is moved back from the resonator to a distance of 2 metres, say, the sparking ceases, and cannot be obtained even on screwing the knobs very near each other, but when the screen is moved twice as far away (4 metres), sparking is again obtained which is more vigorous than when the screen is entirely removed. Without the screen sparking is obtained at every distance, diminishing only in intensity as the distance between the oscillation and resonating circuits is increased. With the screen, on the other hand, sparking is obtained at some places and no sparking at others, even though these may be nearer the vibrator.

It should be observed that no sparking is obtained in the resonator when it is placed immediately behind the zinc sheet, and this shows that the zinc sheet screens off the electromagnetic action of the vibrator.

The interpretation of these results is that waves of electromagnetic disturbance are radiated from the vibrator which fall upon the metal screen and are reflected there. The reflected waves interfere with the direct waves, as in the case of sound. At the distance of a quarter wave (2 metres in the above experiment), or any odd multiple of a quarter wave from the screen, the two are in opposite phases, and their combined effect on the resonator is zero, while at distances of 2, 4, 6, etc., quarter waves from the screen the direct and reflected waves are in accordance, and vigorous sparking is produced in the resonating circuit. The corresponding interference of direct and reflected sound waves is well exhibited in the nodes and loops of organ pipes.

Having once obtained these electromagnetic waves, and the means of observing them, Hertz proceeded to inquire into their laws of reflection and refraction. To this effect he modified the apparatus so as to render it more sensitive, and concentrated the radiation by means of zinc reflectors bent into the shape of parabolic cylinders. The vibrator consists of two brass cylinders (Fig. 242), each about 12 or 13 centimetres long, 3 centimetres in diameter, and rounded or terminated by knobs at the sparking ends. The cylinders are placed in the focal line of the parabolic reflector, and are connected to the terminals of an induction coil which excites the oscillations. The receiver consists of two pieces of thick wire each about 50 centimetres long. They are also placed in the focal line of a parabolic reflector, and from the end of each a thin wire passes at right angles through the reflector to a sparking space at the back, where the induced sparking can be observed without obstructing the waves falling on the reflector. The total length of the vibrator is nearly half that of the emitted wave. For success it is necessary that the pole surfaces of the sparking space should be frequently polished and guarded against



Receiver. Fig. 242. Vibrator.

the illumination of simultaneous lateral discharges, for the rays of higher refrangibility are detrimental to the working of the apparatus.

With this apparatus Hertz detected reflection from the walls and obstacles around the room, and subjected the electromagnetic radiations to most of the experiments which it is customary to perform with light and heat rays. He had a large prism of pitch constructed, and verified that these waves are refracted in passing through it, just as light and heat waves are refracted in passing into a new medium. The angle of the prism was 30°, and a deviation of 22° was observed in the transmitted waves, giving a refractive index of 1·69 for these long electromagnetic waves. The refractive index of pitch for light waves varies from 1·5 to 1·6, but a disagreement in refractive index is not to be taken as evidence against the similarity of light and electromagnetic waves, for it would be remarkable, and even contrary to expectation, if the refractive index of waves 1 metre in length happened to be the same as that of waves 200,000 times shorter.

From the mode of production of these waves it is clear that they consist of transverse vibrations and are plane-polarised, for the electric force is parallel to the vibrator and the lines of magnetic force are circles round it. The directions of the electric and magnetic forces are consequently in the wave front and at right angles to each other. If the receiving mirror be rotated round the direction of the rays coming from the vibrator the sparking gradually ceases, and when the focal line of the receiver is horizontal—that is, at right angles to the focal line of the vibrator—no sparks can be obtained, even though very close to the vibrator. The two pieces of apparatus thus play the part of the polariser and analyser in a polariscope.

The polarisation of the waves may be shown by introducing between the mirrors a wire screen consisting of wires wound parallel to a given direction on a light frame, so as to resemble a diffraction grating. When the direction of the wires is parallel to the oscillator—that is, to the focal lines of the mirrors—the screen is opaque to the radiation, it refuses to transmit the waves but reflects them copiously. On the other hand, when the direction of the wires is perpendicular to the focal lines of the mirrors, the waves are freely transmitted and sparking is obtained in the receiving circuit. The screen is thus opaque to the radiation when the direction of the wires is parallel to the electric force, but transparent when the wires are parallel to the magnetic force. It thus plays the part of an analyser.

Hertz's experiments have been repeated in this country with various modifications, and particularly in Dublin by Professor G. F. FitzGerald, and Dr. F. T. Trouton,[1] who have further studied the cases of reflection from non-conducting substances, such as glass and paraffin. They have also settled the long-disputed question as to the direction of vibration in relation to the plane of polarisation in plane-polarised light. Fresnel and his followers required the luminous vibration to be at right angles to the plane of polarisation, while, according to MacCullagh and Neumann, it must take place in that plane. On the other hand, Maxwell's theory indicated that something occurs in both planes, a magnetic vibration in one and an electric in the other, or rather a vibration accompanied by magnetic force in one direction and electric force in the perpendicular direction. The theory further indicated that the electric force is perpendicular to, and the magnetic force in, the plane of polarisation. This conclusion has been verified by experiment, for the electromagnetic waves are found not to be reflected at the polarising angle from the surface of a bad conductor, such as a wall, when the electric force is parallel to the plane of incidence, but reflection occurs at all angles when the electric force is perpendicular to the plane of incidence. In the polarised ray therefore

[1] F. T. Trouton, *Nature*, 22nd August 1889, etc.

the electric force is perpendicular to, and the magnetic force in, the plane of polarisation.

Hertz's experiments have been repeated by Professor O. J. Lodge and Mr. E. J. Dragoumis[1] at University College, Liverpool, with a slightly modified form of receiver. In order to render the effect of the electric oscillations more easily observable in the resonating circuit, a Geissler's tube was placed in the spark-gap, one electrode of the tube being connected with either side of the gap. When the apparatus is at work the tube lights up and renders the effect of the electric oscillations visible even at a considerable distance. The detection of the oscillations by the chemical action of the current on iodide of potash paper was also tried.

The sympathetic oscillations in the receiver may also be detected by means of a delicate galvanometer. One terminal of the galvanometer is connected to one side of the spark-gap, and the other to the other side of the gap. When sparking occurs at the gap a deflection of the galvanometer occurs, and with a delicate long-coil galvanometer a very marked effect may be produced. Great delicacy of adjustment at the gap is necessary, but when properly arranged the apparatus is very sensitive. By this method Professor FitzGerald[2] has exhibited the effect to a large audience.

The Hertzian vibrations may also be detected by the method of the Bolometer—that is, by the variation, when heated, of the resistance of a thin wire placed across the spark-gap.

**331. Radiation of Electromagnetic Waves.**—When an electric oscillator is at work a periodic variation, or vibration, is occurring at each point of the field around it. The distributions of electric force in the field at the stages $t=\frac{1}{8}T$, $t=\frac{1}{4}T$, $t=\frac{3}{8}T$, $t=\frac{1}{2}T$, have been plotted by Hertz and are represented in Figs. 243-246. If T be the time of a complete vibration, the state of the field at any instant during the first half period is just reversed at the corresponding instant of the second half, so that it is only necessary to trace the changes in the field during half a complete vibration. Fig. 243 shows the field at the time $t=0$, or $t=n\frac{1}{2}T$. Here the poles of the oscillator are free from electrification, and no lines of force run out from them. Electric charges are just about to accumulate on the extremities. Lines of force are about to spring out, and will continue to spring out

[1] E. J. Dragoumis, *Nature*, 4th April 1889. An account of the recent work in this department has been given by Professor O. J. Lodge (the work of Hertz and some of his successors, being an evening lecture at the Royal Institution, 1st June 1894. Reprinted from *The Electrician* by the Electrician Printing and Publishing Company, London).

[2] G. F. FitzGerald, Evening Lecture at the Royal Institution, 21st March 1890.

till $t=\frac{1}{8}$T. At this time the field of force will be as represented in Fig. 244. Here it may be observed the lines of force are not drawn right up to the poles of the oscillator, for the formulæ (Art. 335) from which the curves were plotted regard the oscillator as exceedingly small, so that in the immediate neighbourhood of a finite vibrator they do not sufficiently represent the true state of the field. A small spherical space around the oscillator, to which the formulæ cannot be applied, is consequently left unplotted. The speed at which the disturbance travels out in the initial stage is much greater than the

Fig. 243, $t=0$.

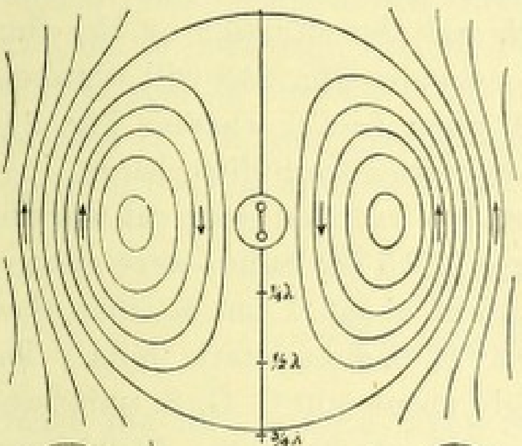
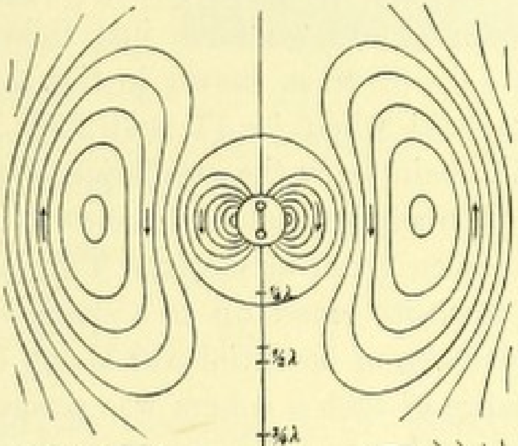



Fig. 244, $t=\frac{1}{8}$T.



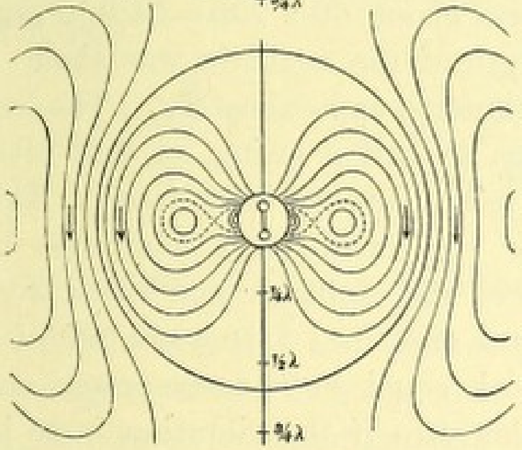




Fig. 246, $t=\frac{3}{8}$T.

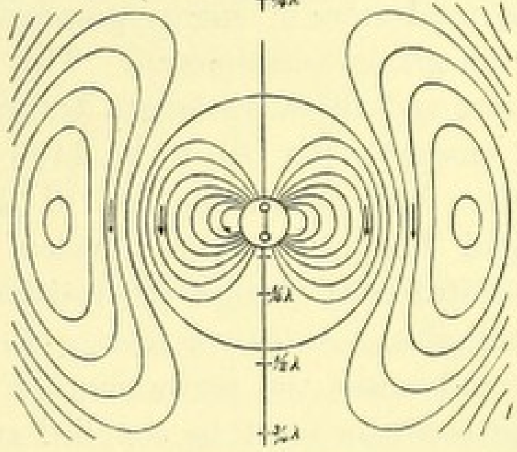



Fig. 245, $t=\frac{2}{8}$T.

normal velocity, and when $t=\frac{1}{4}$T the electric wave has nearly traversed half a wave length instead of a quarter. There is thus a gain of a quarter period, and the velocity finally subsides to the normal value $(1/\sqrt{K\mu})$ at some distance from the origin. Fig. 245 represents the field at the time $t=\frac{1}{4}$T, and Fig 246 that when $t=\frac{3}{8}$T. In the latter a singular action is shown as coming into operation. The lines farthest from the origin are being drawn together by the stress with a lateral inflection. This inflection approaches nearer and nearer to the axis of $z$ till finally a self-closed surface like a vortex ring detaches itself from each of the outer lines, and these spread out into space while the residue sink back into the conductor. The number of receding lines is exactly the same as the number of originally expanding lines, but their energy is diminished by that of the detached portions. These portions carry energy out into space, and in consequence the oscillation must soon subside unless the source be supplied with energy.

Returning now to Fig. 243 it will represent the state of things at the time $t = \frac{1}{2}T$, the closed surfaces being the detached lines of force travelling out into space. The oscillator is again without charge, but charging is about to commence and new lines of force are about to spring out from the poles. These will compress the detached lines whose history we have just followed, and the situation at the time $t = \frac{5}{8}T$ will be as shown in Fig. 244. So also Fig. 245 will represent the field when $t = \frac{3}{4}T$, and Fig. 246 when $t = \frac{7}{8}T$, whereas Fig. 243 will again show the oscillation about to start a fresh disturbance.

An approximate estimate of the energy radiated in actual working has been given by Hertz. Two spheres of 15 cm. radius were charged in opposite senses up to a spark length of about 1 cm., so that their difference of potential was about 120 C. G. S. electrostatic units. The charge of such a sphere was accordingly about 900 C. G. S. units, and the total store of energy possessed was about $60 \times 900 = 54{,}000$ ergs or 55 gramme-centimetres. The length of the oscillator was about 1 metre and the wave length 480 centimetres approximately. The loss of energy comes out to be 2400 ergs to half a swing, so that after eleven half-swings one half of the initial energy is radiated. This rapid damping is made evident by experiment, and could not be obviated even though the resistance of the oscillator and spark were negligible. A loss of energy of 2400 ergs in 1·5/100,000,000 of a second means the performance of work equal to 22 horse-power, and the oscillator must be supplied at this rate if the vibration is to be maintained at a constant intensity, in spite of the radiation. During the first few swings of the vibrator the intensity of the radiation at about 12 metres distant from the vibrator corresponds to the intensity of the solar radiation at the surface of the earth (see p. 569).

Before entering into the mathematical investigation of the properties of these curves, we shall, for convenience, first summarise Maxwell's equations of the electromagnetic field. For this purpose it will be useful, in the first place, to obtain an expression for the line integral of a vector taken round a small rectangular circuit.

**332. Line Integral of a Vector.**—In a field of force the force at any point may in general be represented in magnitude and direction by a right line. Such a quantity, possessing both magnitude and direction, is termed a directed quantity or a *vector*. The line integral of a vector round any curve or circuit may be defined as $\int F \cos \phi ds$ where F represents the value of the vector at any point of the curve, $ds$ the corresponding element of the curve, and $\phi$ the angle between the directions of F and $ds$. Since $F \cos \phi$ is the component of F parallel to $ds$, it is clear that when F is a force, then the line integral is simply the work done in traversing the circuit.

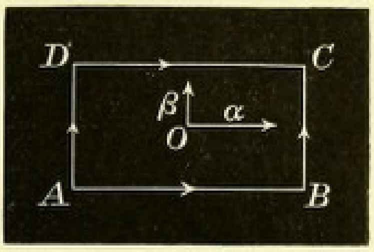


Fig. 247.

Let us now consider a small rectangular circuit of sides $\delta x$ and $\delta y$, and centre O (Fig. 247). Then if the components of F at O be $\alpha$ and $\beta$ respectively in the directions of the sides of the rectangle, and if F varies from point to point of the field, it follows that when the sides of the rectangle are very small the vector components which appear in the line integral round the circuit ABCD will be

$$\alpha - \tfrac{1}{2}\frac{d\alpha}{dy}\delta y \text{ along AB}, \ \beta + \tfrac{1}{2}\frac{d\beta}{dx}\delta x \text{ along BC},$$
$$\alpha + \tfrac{1}{2}\frac{d\alpha}{dy}\delta y \text{ along DC}, \ \beta - \tfrac{1}{2}\frac{d\beta}{dx}\delta x \text{ along AD}.$$

Hence the line integral taken round the rectangle in the direction ABCD is

$$\left(\alpha - \tfrac{1}{2}\frac{d\alpha}{dy}\delta y\right)\delta x + \left(\beta + \tfrac{1}{2}\frac{d\beta}{dx}\delta x\right)\delta y - \left(\alpha + \tfrac{1}{2}\frac{d\alpha}{dy}\delta y\right)\delta x - \left(\beta - \tfrac{1}{2}\frac{d\beta}{dx}\delta x\right)\delta y$$
$$= \left(\frac{d\beta}{dx} - \frac{d\alpha}{dy}\right)\delta x \delta y.$$

The quantity $\frac{d\beta}{dx} - \frac{d\alpha}{dy}$ is termed the *curl* of the vector, and consequently we have found that the line integral of a vector taken round a small rectangular circuit is the product of the area of the circuit by the curl of the vector, or more generally, the line integral of a vector taken round the perimeter of any area is equal to the surface integral of the curl taken over the area.

**333. Fundamental Electromagnetic Equations—Isotropic Dielectric.**—We are now in a position to write down the equations connecting the electric and magnetic forces in any medium. In the first place, for simplicity, let the medium be an isotropic non-conductor such as air or glass.

Then if we consider any circuit placed in such a medium, it is found by experiment that an induced current is generated in the circuit when the magnetic field varies, and that the electromotive force of this current is measured by the rate at which the whole flux of magnetic induction passing through the circuit changes. That is, the line integral of the electric force taken round the circuit is equal to the time rate of change of the surface integral of the magnetic induction passing through it. Hence if we consider elementary rectangular circuits of sides, $\delta y$, $\delta z$; $\delta z$, $\delta x$; $\delta x$, $\delta y$ respectively, having their planes parallel to the planes of reference, then denoting the components of the electric force at the centre of the parallelopiped $\delta x$, $\delta y$, $\delta z$ by P, Q, R, and the components of the magnetic force by $\alpha$, $\beta$, $\gamma$ respectively, we have the line integral of the electric force round the rectangle $\delta y$, $\delta z$ equal to

$$\left(\frac{dR}{dy}-\frac{dQ}{dz}\right)\delta y\delta z,$$

and if the magnetic permeability of the medium be $\mu$, the flux of magnetic induction through the same rectangle is

$$\mu\alpha\delta y\delta z;$$

hence paying attention to sign we have at once the relation

$$-\mu\frac{d\alpha}{dt}=\frac{dR}{dy}-\frac{dQ}{dz} \qquad (1),$$

similarly

$$-\mu\frac{d\beta}{dt}=\frac{dP}{dz}-\frac{dR}{dx} \qquad (2),$$

$$-\mu\frac{d\gamma}{dt}=\frac{dQ}{dx}-\frac{dP}{dy} \qquad (3).$$

That is, the curl of the electric force is equal to the time rate of change of the magnetic induction. In the same way the line integral of magnetic force taken round a circuit is equal to the time rate of change of the flux of electric displacement passing through the circuit, so that if the specific inductive capacity of the medium be K we have the corresponding set of equations [1]

$$K\frac{dP}{dt}=\frac{d\gamma}{dy}-\frac{d\beta}{dz} \qquad (4),$$

$$K\frac{dQ}{dt}=\frac{d\alpha}{dz}-\frac{d\gamma}{dx} \qquad (5),$$

$$K\frac{dR}{dt}=\frac{d\beta}{dx}-\frac{d\alpha}{dy} \qquad (6).$$

[1] The electric current in a straight wire is related to the magnetic force in the field around it by the fact that the line integral of the magnetic force taken round a curve enclosing the wire is equal to the current multiplied by $4\pi$. Thus if the current strength be C the magnetic force at a distance $r$ from the wire is $2C/r$, and the integral of this taken around a circle of radius $r$ is $2\pi r \times 2C/r = 4\pi C$. To apply this to a dielectric, Maxwell introduced the idea of what he termed "displacement" currents. Thus if we consider an electrified conductor placed in a dielectric, the electric density (or displacement) at any point of the conductor being $\rho$, the electric

We have thus six equations connecting the six quantities, $\alpha$, $\beta$, $\gamma$, P, Q, R, and from these we may easily deduce a differential equation for any one of them. Thus if we differentiate (1) with respect to $t$ and substitute from (5) and (6) for $dQ/dt$ and $dR/dt$ we have

$$\mu K\frac{d^2\alpha}{dt^2}=\frac{d}{dy}\left(\frac{d\alpha}{dy}-\frac{d\beta}{dx}\right)-\frac{d}{dz}\left(\frac{d\gamma}{dx}-\frac{d\alpha}{dz}\right)$$
$$=\frac{d^2\alpha}{dx^2}+\frac{d^2\alpha}{dy^2}+\frac{d^2\alpha}{dz^2}-\frac{d}{dx}\left(\frac{d\alpha}{dx}+\frac{d\beta}{dy}+\frac{d\gamma}{dz}\right).$$

But since we have supposed the medium to be an isotropic non-conductor in which the magnetic density is zero; or, in other words, that it contains no source or sink of magnetic force, it follows that the flux

force just outside the surface of the conductor will be $4\pi\rho$ in air and $4\pi\rho/K$ in any other medium. This displacement is supposed to exist wherever there is electric force, and a variation of the displacement corresponds to an electric current, the current strength being $d\rho/dt$. As the displacement at any point of an isotropic dielectric is a directed quantity, being in the direction of the electric force, it follows that if its components parallel to the axes be denoted by $f$, $g$, $h$ at any point, then the components of the force at that point are

$$P=4\pi f/K, \quad Q=4\pi g/K, \quad R=4\pi h/K,$$

consequently the components of the displacement current are

$$u=\frac{df}{dt}=\frac{K}{4\pi}\frac{dP}{dt}, \quad v=\frac{dg}{dt}=\frac{K}{4\pi}\frac{dQ}{dt}, \quad w=\frac{dh}{dt}=\frac{K}{4\pi}\frac{dR}{dt}.$$

Hence if we apply the above law connecting the current strength with the line integral of magnetic force we have

$$\frac{d\gamma}{dy}-\frac{d\beta}{dz}=4\pi u=K\frac{dP}{dt}.$$

Thus the curl of the magnetic force is $4\pi$ times the electric current, and, expressing equations 1, 2, 3 in similar language, we may say that the curl of the electric force is $4\pi$ times the "magnetic current." Thus KP, KQ, KR are the components of what may be called the electric flux density, and $KdP/dt$, $KdQ/dt$, $KdR/dt$ are the components of the electric current density. Similarly $\mu\alpha$, $\mu\beta$, $\mu\gamma$ are the components of the magnetic flux density, and $\mu d\alpha/dt$, $\mu d\beta/dt$, $\mu d\gamma/dt$ the components of the magnetic current density. When the medium is not a perfect non-conductor the components of the conduction current must be taken into consideration, and if the electric conductivity be $k$ these components will be $kP$, $kQ$, $kR$, so that the equations become

$$\left(4\pi k+K\frac{d}{dt}\right)P=\frac{d\gamma}{dy}-\frac{d\beta}{dz}, \text{ etc.}$$
$$-\left(4\pi g+\mu\frac{d}{dt}\right)\alpha=\frac{dR}{dy}-\frac{dQ}{dz}, \text{ etc.}$$

The factor $g$ which appears in the latter equation has been introduced by Mr. Oliver Heaviside, and is the magnetic analogue of $k$—that is, it is the magnetic conductivity or the reciprocal of the magnetic resistivity.

of force taken over the surface of an element of volume is zero—that is,

$$\frac{d\alpha}{dx}+\frac{d\beta}{dy}+\frac{d\gamma}{dz}=0.$$

Consequently if we, as usual, denote the operator $\frac{d^2}{dx^2}+\frac{d^2}{dy^2}+\frac{d^2}{dz^2}$ by $\nabla^2$ the equation for $\alpha$ becomes

$$\mu\mathrm{K}\frac{d^2\alpha}{dt^2}=\nabla^2\alpha,$$

with similar equations for $\beta$ and $\gamma$.

In the same way by differentiating (4) with regard to $t$ and substituting from (2) and (3) we obtain the equation

$$\mu\mathrm{K}\frac{d^2\mathrm{P}}{dt^2}=\nabla^2\mathrm{P}-\frac{d}{dx}\left(\frac{d\mathrm{P}}{dx}+\frac{d\mathrm{Q}}{dy}+\frac{d\mathrm{R}}{dz}\right).$$

But since there is no electricity distributed through the dielectric the electric density is zero at every point of it; in other words, there is no source or sink of electric force, and therefore

$$\frac{d\mathrm{P}}{dx}+\frac{d\mathrm{Q}}{dy}+\frac{d\mathrm{R}}{dz}=0,$$

so that the differential equations for the electric force become

$$\mu\mathrm{K}\frac{d^2\mathrm{P}}{dt^2}=\nabla^2\mathrm{P},\quad \mu\mathrm{K}\frac{d^2\mathrm{Q}}{dt^2}=\nabla^2\mathrm{Q},\quad \mu\mathrm{K}\frac{d^2\mathrm{R}}{dt^2}=\nabla^2\mathrm{R}.$$

The components of the electric and magnetic forces thus satisfy the same differential equation, and this equation is of the same form as that which determines the vibration of an elastic solid. It follows, therefore, that a periodic electromagnetic disturbance is propagated after the manner of a wave motion in an elastic solid, and that the velocity of propagation is determined by the equation

$$v=\frac{1}{\sqrt{\mu\mathrm{K}}}.$$

**334. Plane Wave in an Isotropic Non-conductor.**—Let us now consider the simple case of a plane wave propagated through an isotropic dielectric. Let the plane of the wave front be parallel to the plane $xy$, so that the axis of $z$ is normal to the wave front—that is, parallel to the direction of propagation. In this case all the quantities which determine the character of the wave are functions of $z$ and $t$, being independent of $x$ and $y$. Hence the equations (3) and (6) of Art. 333 become

$$\mu\frac{d\gamma}{dt}=0,\quad \text{and } \mathrm{K}\frac{d\mathrm{R}}{dt}=0,$$

which show that $\gamma$ and R are each independent of $t$—that is, they are each zero (or constant) and play no part in the propagation of the wave. We conclude, therefore, that the electric and magnetic forces have no component (or at least no periodic component) normal to the wave front—that is, the disturbance is confined to the plane of the wave front or is transverse. An electromagnetic wave consequently possesses the essential property of a light wave in having the vibration in the wave front, or transverse to the direction of propagation.

Further, it is clear that if the axes of $x$ and $y$ be so chosen that the axis of $x$ is parallel to the direction of the resultant magnetic force in the wave front, then $\alpha$ will represent the whole magnetic force, so that we have $\beta = 0$ permanently, and consequently by equation (2), since R is always zero, it follows that P is always zero. From this it follows that Q must be the resultant electric force—that is, if the resultant magnetic force is parallel to the axis of $x$, then the resultant electric force is parallel to the axis of $y$. In other words, the electric and magnetic forces are in the wave front, and are at right angles to each other.

We see, therefore, that according to the electromagnetic theory of light the vibration is a transverse periodic disturbance attended by electric force in one direction, and magnetic force in the perpendicular direction. The former is the direction of vibration postulated by the theory of Fresnel, and the latter that demanded by MacCullagh. The war between the rival theories is consequently at an end, for one speaks of the direction of the electric force while the other speaks of the magnetic force, so that both are equally right and equally wrong in dealing with the vibration as a whole.

The general differential equation of propagation becomes simplified in the case of a plane wave travelling in the direction of the axis of $z$. For since P, Q, R, $\alpha$, $\beta$, $\gamma$ are independent of $x$ and $y$, the equations of Art. 333 become

$$\mu K \frac{d^2P}{dt^2} = \frac{d^2P}{dz^2}, \quad \mu K \frac{d^2Q}{dt^2} = \frac{d^2Q}{dz^2}, \quad \frac{d^2R}{dt^2} = 0,$$

with similar equations for $\alpha$, $\beta$, $\gamma$. Hence while R and $\gamma$ are each zero (or constants), the general expressions for P, Q, $\alpha$, $\beta$ are of the form

$$P = \phi(z - vt) + \psi(z + vt),$$

where $v$ is the velocity of propagation and is equal to $1/\sqrt{\mu K}$. In the case of air K is equal to unity if the electrostatic system of units be employed, and $\sqrt{\mu}$ is the reciprocal of a velocity; but if the electromagnetic system be used $\mu = 1$ and $\sqrt{K}$ is the reciprocal of a velocity. Hence if the light vibration be an electromagnetic disturbance, this velocity should be equal to that of light—that is, of light waves of the length here considered—and this conclusion is supported by the results of observation. Again, since the refractive index is inversely as the wave velocity, it follows that if we take the magnetic permeability equal to unity we should have the refractive index proportional to the square root of the specific inductive capacity. As the refractive index varies with the wave length, the limiting value of the refractive index must be used if we deal with the very long waves considered in the measurement of specific inductive capacity.

**335. The Hertzian Oscillator.**—In order to investigate the field in the neighbourhood of a simple dumb-bell oscillator Professor Hertz[1] employed the general equations

$$\left.\begin{aligned} -\mu\frac{d\alpha}{dt}&=\frac{dR}{dy}-\frac{dQ}{dz}\\ -\mu\frac{d\beta}{dt}&=\frac{dP}{dz}-\frac{dR}{dx}\\ -\mu\frac{d\gamma}{dt}&=\frac{dQ}{dx}-\frac{dP}{dy}\end{aligned}\right\}(1)\qquad \left.\begin{aligned} K\frac{dP}{dt}&=\frac{d\gamma}{dy}-\frac{d\beta}{dz}\\ K\frac{dQ}{dt}&=\frac{d\alpha}{dz}-\frac{d\gamma}{dx}\\ K\frac{dR}{dt}&=\frac{d\beta}{dx}-\frac{d\alpha}{dy}\end{aligned}\right\}(2),$$

together with the constant flux conditions

$$\frac{d\alpha}{dx}+\frac{d\beta}{dy}+\frac{d\gamma}{dz}=0,\quad\text{and}\ \frac{dP}{dx}+\frac{dQ}{dy}+\frac{dR}{dz}=0.$$

Now if the axis of $z$ be taken to coincide with the axis of the oscillator (Fig. 248), the lines of magnetic force will be circles having their centres on the axis of $z$, and their planes perpendicular to it. Hence we have $\gamma=0$, and the flux condition becomes

$$\frac{d\alpha}{dx}+\frac{d\beta}{dy}=0.$$

This means that the quantity $\alpha dy-\beta dx$ is a complete differential of some function, and if we use Hertz's notation and represent this function by $d\Pi/dt$, we have

$$\alpha=\frac{d^2\Pi}{dydt},\quad \beta=-\frac{d^2\Pi}{dxdt},\quad \gamma=0.$$

Substituting these values of $\alpha$, $\beta$, $\gamma$ in the equations (2) we find

$$K\frac{dP}{dt}=\frac{d}{dt}\left(\frac{d^2\Pi}{dxdz}\right),\quad K\frac{dQ}{dt}=\frac{d}{dt}\left(\frac{d^2\Pi}{dydz}\right),\quad K\frac{dR}{dt}=-\frac{d}{dt}\left(\frac{d^2\Pi}{dx^2}+\frac{d^2\Pi}{dy^2}\right).$$

---

[1] The theory of the electric oscillator was given by Hertz, together with the experiments already referred to, in *Wied. Ann.*, 1888-89. These researches have been recently translated into English by Mr. D. E. Jones (*Electric Waves*, London, Macmillan and Co., 1893).

Now these equations show that the quantities

$$KP-\frac{d^2\Pi}{dxdz},\quad KQ-\frac{d^2\Pi}{dydz},\quad KR+\left(\frac{d^2\Pi}{dx^2}+\frac{d^2\Pi}{dy^2}\right),$$

are each independent of $t$, so that at any point, as far as a periodic disturbance is concerned, we may take

$$KP=\frac{d^2\Pi}{dxdz},\quad KQ=\frac{d^2\Pi}{dydz},\quad KR=-\left(\frac{d^2\Pi}{dx^2}+\frac{d^2\Pi}{dy^2}\right) \qquad (3),$$

and these values of P, Q, R will satisfy the equations (1) provided $\Pi$ satisfies the equation [1]

$$\mu K\frac{d^2\Pi}{dt^2}=\nabla^2\Pi,$$

for on substitution we find all the equations satisfied identically. It may be remarked that the equation determining $\Pi$ is the same in form as that satisfied by the electric and magnetic forces.

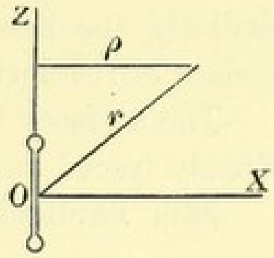


Fig. 248.

When the distribution is symmetrical round the axis of $z$, as in the case of a simple electric oscillator (Fig. 248), the forces at any point of the field will depend only on the $z$ co-ordinate of the point and on the distance $\rho=\sqrt{x^2+y^2}$ from the axis of $z$. Denoting the electric force in the direction of $\rho$ (that is, perpendicular to $z$) by E, and the magnetic force perpendicular to the meridian plane drawn through the oscillator (which in this case is the direction of the resultant magnetic force) by H, we have

$$E=P\frac{x}{\rho}+Q\frac{y}{\rho},\quad H=\alpha\frac{y}{\rho}-\beta\frac{x}{\rho},$$

and by the foregoing equations (3) we have

$$KE=\frac{1}{\rho}\frac{dV}{dz},\quad H=\frac{1}{\rho}\frac{dV}{dt},\quad KR=-\frac{1}{\rho}\frac{dV}{d\rho} \qquad (4),$$

where

$$V=\rho\frac{d\Pi}{d\rho}.$$

The function V holds an important place in the investigation of the field. We shall show that the curves in which the surfaces of revolution V = const. are cut by the meridian planes are the lines of electric force. For if the direction of the resultant electric force at any point

[1] The condition required is that $\mu K\frac{d^2\Pi}{dt^2}-\nabla^2\Pi$ shall be independent of $x$ and $y$—that is, at any point it may be a function of $z$ and $t$; but since the electric and magnetic forces considered above involve differentiation of $\Pi$ with regard to $x$ and $y$, this function of $z$ and $t$ may be taken as zero without affecting the field.

makes an angle $\phi$ with the axis of $z$ we have

$$\tan\phi = \mathrm{E}/\mathrm{R} = -\frac{d\mathrm{V}}{dz}\Big/\frac{d\mathrm{V}}{d\rho}$$

by (4), which is the trigonometrical tangent of the angle which the tangent line to the meridian section of the surface V = const. makes with the axis of $z$. Hence the resultant electric force is tangential to the surface V, and along its curve of section with the meridian plane.

Again, if two surfaces V and $\mathrm{V} + d\mathrm{V}$ be taken enclosing a shell, and if any plane perpendicular to the axis of $z$ be drawn cutting the shell, the flow of electric force across the strip of the plane intercepted by the shell will be constant. For the area of the strip is $2\pi\rho d\rho$ and the force perpendicular to it is R, therefore the flow is

$$\int \mathrm{KR} \,.\, 2\pi\rho d\rho = -\int 2\pi \frac{d\mathrm{V}}{d\rho}\, d\rho = 2\pi(\mathrm{V}_1 - \mathrm{V}_2),$$

similarly the flow across the strip intercepted by the shell on any surface of revolution round the axis of $z$ is constant.

The surfaces V are those whose history and development we have already traced in Figs. 243-246.

**336. Particular Solution.**—A particular solution of the equation

$$\mu\mathrm{K}\frac{d^2\Pi}{dt^2} = \nabla^2\Pi$$

is

$$\Pi = \frac{el}{r}\sin(mr - nt) \qquad (1),$$

where[1] $e$, $l$, $m$, $n$ are constant quantities and $m$ and $n$ are taken to satisfy the relation

$$\frac{m}{n} = \sqrt{\mu\mathrm{K}}.$$

that is, $m = 2\pi/\lambda$, $n = 2\pi/\mathrm{T}$, and $n/m$ is the velocity of propagation.

The function V is now easily determined, for by equation (1) we have

$$\frac{d\Pi}{d\rho} = \frac{elm}{r}\left\{\cos(mr - nt) - \frac{\sin(mr - nt)}{mr}\right\}\frac{dr}{d\rho},$$

but $dr/d\rho = \sin\theta$, and $\rho/r = \sin\theta$, therefore

$$\mathrm{V} = \rho\frac{d\Pi}{d\rho} = elm\left\{\cos(mr - nt) - \frac{\sin(mr - nt)}{mr}\right\}\sin^2\theta \qquad (2).$$

From this value of V we may at once obtain the forces by differentiation. We shall first examine the case of any point situated in the

[1] Hertz writes $el$ instead of a single constant A, the signification being, as will be seen afterwards, that $l$ is the length of the oscillator and $e$ the charge on either pole.

equatorial plane $xy$—that is, the plane through the middle point of the oscillator perpendicular to its length.

**337. Equatorial Plane.**—In the equatorial plane, $xy$, we have

$$\theta=90^\circ, \quad dz=-rd\theta, \quad \rho=r, \quad d\rho=dr.$$

Therefore by substituting in equation (2) of the foregoing article we obtain

$$\mathrm{KE}=\frac{1}{\rho}\frac{d\mathrm{V}}{dz}=-\frac{1}{r^2}\frac{d\mathrm{V}}{d\theta}=0,$$

$$\mathrm{H}=\frac{1}{\rho}\frac{d\mathrm{V}}{dt}=\frac{elmn}{r}\left\{\sin(mr-nt)+\frac{\cos(mr-nt)}{mr}\right\},$$

$$\mathrm{KR}=\frac{1}{\rho}\frac{d\mathrm{V}}{d\rho}=\frac{elm^2}{r}\left\{-\sin(mr-nt)-\frac{\cos(mr-nt)}{mr}+\frac{\sin(mr-nt)}{m^2r^2}\right\}.$$

The resultant electric force is therefore perpendicular to the equatorial plane, for since $\mathrm{E}=0$ the resultant force is R, and is parallel to the vibrator.

Writing the expression for R in the form

$$\mathrm{KR}=\mathrm{A}\sin(nt-\delta),$$

we have at once by comparison

$$-\mathrm{A}\sin\delta=\frac{el}{r^3}(m^2r^2\sin mr+mr\cos mr-\sin mr),$$

$$-\mathrm{A}\cos\delta=\frac{el}{r^3}(m^2r^2\cos mr-mr\sin mr-\cos mr).$$

Hence by squaring and adding we obtain

$$\mathrm{A}=\frac{el}{r^3}\sqrt{1-m^2r^2+m^4r^4}.$$

At small distances from the origin the amplitude A varies inversely as the cube of $r$, but for great values of $r$ the amplitude approximates to the inverse ratio of the distance.

To determine $\delta$ we have

$$\tan\delta=\frac{m^2r^2\sin mr+mr\cos mr-\sin mr}{m^2r^2\cos mr-mr\sin mr-\cos mr}=\frac{\tan mr-\dfrac{mr}{1-m^2r^2}}{1+\dfrac{mr}{1-m^2r^2}\tan mr}.$$

Therefore

$$\delta=mr-\tan^{-1}\frac{mr}{1-m^2r^2},$$

and we have finally

$$\mathrm{KR}=\frac{el}{r^3}\sqrt{1-m^2r^2+m^4r^4}\sin\left(nt-mr+\tan^{-1}\frac{mr}{1-m^2r^2}\right).$$

The expression for $\delta$ shows that the phase at a given instant is not simply proportional to the distance of the point under consideration from the oscillator, but is less by the angle whose tangent is

$mr/(1 - m^2r^2)$. When $r$ is very large this angle is approximately equal to $\pi$, hence as we recede from the oscillator the value of $\delta$ becomes more and more nearly equal to $mr - \pi$. Very near the oscillator we have $\delta = 0$, since $\delta$ and $r$ vanish together. For small values of $r$ it is clear that $\delta$ is negative, so that the curve (Fig. 249) representing the relation between $\delta$ and $r$, at any given instant, lies below the axis OX near the origin, the distance $r$ being measured along OX and the angle $\delta$ along OY. The negative value of $\delta$ will be greatest when $d\delta/dr = 0$, which corresponds to $mr = \sqrt{2}$—that is, $r = \lambda/\pi\ \sqrt{2} = \lambda/4\ .\ 4$ approximately. From this point the curve begins to rise, and as $r$ increases $\delta$ increases almost proportionately, approaching more nearly the value of $mr - \pi$. The curve is consequently asymptotic to the line $\delta = mr - \pi$, which meets the axis OY at a distance numerically equal to $\pi$ below the origin, and is inclined to the axis OX at an angle of which the trigonometrical tangent is numerically equal to $m$.

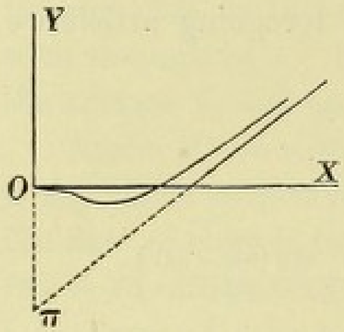


Fig. 249.

The velocity of propagation is determined as the speed ($dr/dt$) with which any surface of equal phase moves forward. Taking this surface to be that in which $R = 0$, or the phase a multiple of $\pi$, we have then

$$nt - mr + \tan^{-1}\frac{mr}{1 - m^2r^2} = 0, \quad \text{or } N\pi.$$

Hence

$$\tan(mr - nt) = \frac{mr}{1 - m^2r^2},$$

a result which follows also directly from the original equation for R. Therefore

$$v = \frac{dr}{dt} = \frac{n}{m} \cdot \frac{1 - m^2r^2 + m^4r^4}{m^2r^2(m^2r^2 - 2)}.$$

This expression shows that the value of $v$ is infinite both at the origin and at the distance $r = \sqrt{2}/m = \lambda/4\ .\ 4$. This is the point where the tangent to the curve (Fig. 249) is parallel to OX, the interpretation being that at this point a small distance is traversed without change of phase. At points between this point and the origin $v$ is negative, and for greater values of $r$ the velocity is positive. Hence the wave spreads out from this point in both directions. It is here that they are thrown off into space, part being radiated and part receding again into the oscillator, as indicated in Hertz's diagrams (Figs. 243-246). For large values of $r$ the velocity approaches more nearly the normal velocity $v = n/m = \lambda/T$.

The magnetic force may be treated in the same manner. Thus, if we write

$$H = A' \sin(nt - \delta'),$$

we have

$$A' \sin \delta' = -\frac{cln}{r^2}(mr \sin mr + \cos mr),$$

$$A' \cos \delta' = \frac{cln}{r^2}(\sin mr - mr \cos mr).$$

Hence

$$A' = \frac{cln}{r^2}\sqrt{1 + m^2r^2},$$

and

$$\tan \delta' = \frac{mr \sin mr + \cos mr}{mr \cos mr - \sin mr} = \frac{\tan mr + \dfrac{1}{mr}}{1 - \dfrac{1}{mr}\tan mr}$$

or

$$\delta' = mr + \tan^{-1}\frac{1}{mr},$$

and we have finally

$$H = \frac{cln}{r^2}\sqrt{1 + m^2r^2}\sin\left(nt - mr - \tan^{-1}\frac{1}{mr}\right).$$

At the origin we have $r = 0$, and therefore $\delta' = \frac{1}{2}\pi$, so that initially the electric and magnetic components differ in phase by a quarter period. At great distances, however, we have $\delta' = mr$, so that the phase of the magnetic component does not increase proportionately to $r$, but there is a loss of $\frac{1}{2}\pi$, with the result that at a distance from the origin the electric and magnetic components differ in phase by $\pi$.

The curve (Fig. 250) indicating the relation between $\delta'$ and $r$ will therefore meet the axis OY at a distance numerically equal to $\frac{1}{2}\pi$, and will finally approach asymptotically the line $\delta = mr$. This line passes through the origin and is inclined to the axis OX at an angle whose tangent is numerically equal to $m$. This curve and that of Fig. 249 are therefore asymptotic to lines which are parallel, and the distance between these lines, measured parallel to OY, is numerically equal to $\pi$, showing that the phases of the electric and magnetic components differ by $\pi$.

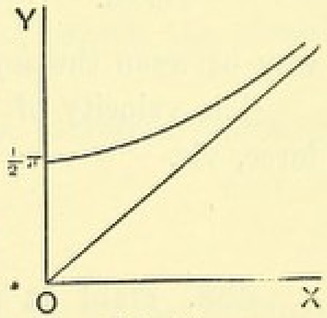


Fig. 250.

The velocity of propagation of the magnetic component may be found by considering the wave front where the magnetic force is zero. We have then

$$\tan(nt - mr) = \frac{1}{mr}.$$

Therefore

$$v = \frac{dr}{dt} = \frac{n}{m} \cdot \frac{1 + m^2r^2}{m^2r^2}.$$

When $r=0$—that is, at the vibrator—the velocity is infinite, but on receding from the vibrator the velocity rapidly approximates to the normal value $n/m=\lambda/T$.

For the time rate of change of the magnetic force we have

$$\frac{d\mathrm{H}}{dt}=\frac{elmn^2}{r}\left\{\frac{\sin(mr-nt)}{mr}-\cos(mr-nt)\right\}.$$

Throwing this expression into the form

$$\mathrm{A}''\sin(nt-\delta''),$$

we find at once

$$\mathrm{A}''\sin\delta''=\frac{eln^2}{r^2}(mr\cos mr-\sin mr),$$

$$\mathrm{A}''\cos\delta''=-\frac{eln^2}{r^2}(mr\sin mr+\cos mr).$$

Therefore

$$\mathrm{A}''=\frac{eln^2}{r^2}\sqrt{1+m^2r^2},$$

$$\tan\delta''=\frac{\sin mr-mr\cos mr}{\cos mr+mr\sin mr}=\frac{\tan mr-mr}{1+mr\tan mr},$$

or

$$\delta''=mr-\tan^{-1}mr.$$

At the origin $\delta''$ is zero, and at great distances $\delta''=mr-\frac{\pi}{2}$, so that the curve representing the variation of $\delta''$ with $r$ passes through the origin and finally approaches asymptotically the line $\delta''=mr-\frac{\pi}{2}$, which meets the axis OX at a distance $\frac{\pi}{2}$ below the origin, and the tangent of whose inclination to the axis OX is $m$. It consequently lies halfway between the asymptotic lines of Figs. 249 and 250.

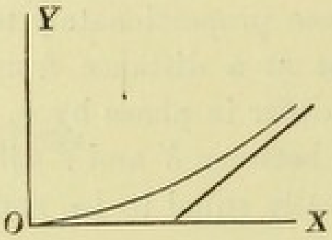


Fig. 251.

The velocity of propagation is the same as that of the magnetic force, viz.

$$v=\frac{n}{m}\frac{1+m^2r^2}{m^2r^2}.$$

**338. Field at a Distance from the Vibrator.**—At a considerable distance from the vibrator we may neglect the higher powers of $1/r$, so that we have

$$\mathrm{V}=elm\cos(mr-nt)\sin^2\theta,$$

$$\mathrm{KE}=-\frac{elm^2}{r}\sin(mr-nt)\sin\theta\cos\theta,$$

$$\mathrm{KR}=\frac{elm^2}{r}\sin(mr-nt)\sin^2\theta,$$

$$\mathrm{H}=\frac{elmn}{r}\sin(mr-nt)\sin\theta.$$

From the second and third equations it follows that

$$R\cos\theta + E\sin\theta = 0,$$

which shows that the direction of the resultant electric force is everywhere perpendicular to the radius vector $r$—that is, the propagation takes place in waves of pure transverse vibrations.

The magnitude of the resultant electric force is

$$R\sin\theta - E\cos\theta = \frac{elm^2}{Kr}\sin\theta\sin(mr - nt).$$

Its amplitude is therefore inversely as the distance and directly proportional to the cosine of the angle which the radius $r$ makes with the equatorial plane.

**339. Field near the Vibrator.**—In the immediate vicinity of the vibrator ($r$ very small) we may put $mr$ vanishing in comparison with $nt$, so that the equations of Art. 336 become

$$\Pi = -\frac{el}{r}\sin nt, \quad V = \frac{el}{r}\sin nt\sin^2\theta,$$

and since

$$\left(\frac{d^2}{dx^2} + \frac{d^2}{dy^2}\right)\frac{1}{r} = -\frac{d^2}{dz^2}\left(\frac{1}{r}\right),$$

we have by equations (3) of Art. 335

$$KP = +\frac{d}{dx}\left(\frac{d\Pi}{dz}\right), \quad KQ = +\frac{d}{dy}\left(\frac{d\Pi}{dz}\right), \quad KR = +\frac{d}{dz}\left(\frac{d\Pi}{dz}\right),$$

so that the electric forces appear as derived from a potential function

$$\Psi = -\frac{d\Pi}{dz} = +el\sin nt\frac{d}{dz}\left(\frac{1}{r}\right) = -\frac{el}{r^2}\sin nt\cos\theta.$$

This force distribution will be that due to the action of a very short rectilinear oscillator of length $l$ on the poles of which the electric charges are, at the maximum, equal to $\pm e$.

For the magnetic force perpendicular to the meridian planes we have

$$H = \frac{1}{\rho}\frac{dV}{dt} = \frac{eln}{r^2}\cos nt\sin\theta,$$

and this represents the magnetic action of an electric current of length $l$ along the vibrator, and of which the intensity is $en\cos nt$, due to the oscillation of the charge $e$, the maximum values of the current being $\pm ne$.

In the axis of $z$ we have

$$\theta = 0, \quad dz = dr, \quad d\rho = rd\theta.$$

Hence

$$E = 0, \quad H = 0, \quad KR = \frac{1}{\rho}\frac{dV}{d\rho} = \frac{2elm}{r^2}\left\{\cos(mr - nt) - \frac{\sin(mr - nt)}{mr}\right\}.$$

The electric force is therefore entirely in the direction of the oscillation. It diminishes for small distances as the inverse cube of $r$ and for large distances as the inverse square.

**340. Radiation of Energy.**—In order to calculate the rate at which energy is radiated by the vibrator we have the equations (see Maxwell's *Electricity and Magnetism*)

$$E=\frac{K}{8\pi}\iiint(P^2+Q^2+R^2)dxdydz,$$

and

$$T=\frac{\mu}{8\pi}\iiint(\alpha^2+\beta^2+\gamma^2)dxdydz,$$

where E is the electrostatic energy in the region through which the integration extends, and T the magnetic energy in the same region. These equations merely express that work is half the product of the final stress by the final strain, and in the case of the electric field, where the electric force is P, the electric displacement is $P/4\pi$.

Now if we multiply the equations, 1, 2, 3 of Art. 333 by $\alpha$, $\beta$, $\gamma$ respectively, and the equations 4, 5, 6 by P, Q, R, we have at once

$$4\pi\frac{d}{dt}(E+T)=\iiint\Big\{P\left(\frac{d\gamma}{dy}-\frac{d\beta}{dz}\right)+Q\left(\frac{d\alpha}{dz}-\frac{d\gamma}{dx}\right)+R\left(\frac{d\beta}{dx}-\frac{d\alpha}{dy}\right)$$
$$-\alpha\left(\frac{dR}{dy}-\frac{dQ}{dz}\right)-\beta\left(\frac{dP}{dz}-\frac{dR}{dx}\right)-\gamma\left(\frac{dQ}{dx}-\frac{dP}{dy}\right)\Big\}dxdydz$$
$$=\iiint\left\{\frac{d}{dx}(\beta R-\gamma Q)+\frac{d}{dy}(\gamma P-\alpha R)+\frac{d}{dz}(\alpha Q-\beta P)\right\}dxdydz,$$

and this integral taken throughout any volume may be expressed as a surface integral taken over the enclosing surface in the form

$$4\pi\frac{d}{dt}(E+T)=\iint\left\{(\beta R-\gamma Q)\lambda+(\gamma P-\alpha R)\mu+(\alpha Q-\beta P)\nu\right\}dS,$$

where $dS$ is an element of the surface, and $\lambda$, $\mu$, $\nu$ the direction cosines of the normal to it.

Now in the case of the simple dumb-bell oscillator everything is symmetrical around the axis of $z$, so that in the plane $xz$ at a considerable distance from the oscillator we have

$$\alpha=\gamma=0, \quad \text{and } \beta=\frac{elmn}{r}\sin(mr-nt)\sin\theta,$$

while

$$KP=-\frac{elm^2}{r}\sin(mr-nt)\sin\theta\cos\theta,$$

$$KR=\frac{elm^2}{r}\sin(mr-nt)\sin^2\theta,$$

$$\lambda=\sin\theta, \quad \mu=0, \quad \nu=\cos\theta.$$

Hence if the integration be extended over the surface of a sphere of radius $r$, having its centre at the centre of the oscillator, we may write $dS = 2\pi r \sin\theta \,.\, r d\theta$, so that we have

$$4\pi\frac{d}{dt}(E+T) = \int \beta(R\sin\theta - P\cos\theta)2\pi r^2 \sin\theta d\theta$$

$$= \frac{4\pi e^2 l^2 m^3 n}{K}\sin^2(mr - nt)\int_0^{\frac{\pi}{2}}\sin^3\theta d\theta.$$

Consequently we have

$$\frac{d}{dt}(E+T) = \tfrac{2}{3}\frac{e^2 l^2 m^3 n}{K}\sin^2(mr - nt).$$

Integrating this between the limits T and zero we obtain the expression for the energy which passes across the whole surface of the sphere during a complete oscillation—that is, since $m = 2\pi/\lambda$ and $n = 2\pi/T$,

$$\frac{e^2 l^2 m^3 n}{3K}\int_0^T \{1 - \cos 2(mr - nt)\}dt$$

$$= \frac{e^2 l^2 m^3 n T}{3K} = \frac{(2\pi)^4 e^2 l^2}{3K\lambda^3},$$

so that the energy radiated [1] per second is

$$\frac{(2\pi)^4 e^2 l^2}{3K\lambda^3 T}.$$

In one of the oscillators used by Hertz the conductors were two equal spheres of 15 cms. radius. These were charged in opposite senses up to a sparking distance of about 1 cm. Taking this to represent a difference of potential of 120 C. G. S. electrostatic units, so that one sphere was charged to a potential of $+60$, and the other to $-60$, then the charge of each sphere was $15 \times 60 = 900$ C. G. S. units. Hence the whole stock of energy possessed by the oscillator at the start amounted to $2 \times \frac{1}{2} \times 900 \times 60 = 54{,}000$ ergs, which is about the energy acquired by a mass of 1 gramme in falling through a height of 55 cm. The length $l$ of the oscillator was 100 cm., and the wave length was about 480 cm., so that the energy lost in the first half-period oscillation was about 2400 ergs, and after eleven half-period oscillations about half the total stock of energy was lost.

**341. Crystalline Dielectric.**—When the medium occupying the field is æolotropic, the specific inductive capacity K, and the magnetic permeability $\mu$ may vary from point to point, and they may also be different for different directions around the same point. Now the

[1] The energy radiated by a small circular current varying according to the simple harmonic law was calculated as early as 1883 by Professor G. F. FitzGerald (*Trans. Royal Dublin Society*, March 1884).

effect of a change of K in crossing any surface drawn in a dielectric is the same as if a certain fictitious distribution of electricity existed on that surface, and consequently the variations of $\mu$ and K throughout any medium are equivalent to certain fictitious distributions of magnetism and electricity through the medium supposed otherwise isotropic. For this reason the quantity

$$\frac{dP}{dx}+\frac{dQ}{dy}+\frac{dR}{dz}$$

will not be zero in an unelectrified æolotropic medium.

If we confine our attention to the case of a homogeneous crystalline medium possessing three mutually rectangular axes of electric symmetry, and if we assume that these axes of electric symmetry are also axes of magnetic symmetry, then the equations of Art. 335 may be written in the form

$$\left.\begin{aligned} -\mu_1\frac{d\alpha}{dt}&=\frac{dR}{dy}-\frac{dQ}{dz}\\ -\mu_2\frac{d\beta}{dt}&=\frac{dP}{dz}-\frac{dR}{dx}\\ -\mu_3\frac{d\gamma}{dt}&=\frac{dQ}{dx}-\frac{dP}{dy}\end{aligned}\right\}\ (1),\qquad \left.\begin{aligned} K_1\frac{dP}{dt}&=\frac{d\gamma}{dy}-\frac{d\beta}{dz}\\ K_2\frac{dQ}{dt}&=\frac{d\alpha}{dz}-\frac{d\gamma}{dx}\\ K_3\frac{dR}{dt}&=\frac{d\beta}{dx}-\frac{d\alpha}{dy}\end{aligned}\right\}\ (2),$$

where $K_1$, $K_2$, $K_3$ are the three principal electric inductive capacities, and $\mu_1$, $\mu_2$, $\mu_3$ the three principal magnetic permeabilities.

As a first case, if we take the magnetic permeability to be the same in all directions, so that $\mu_1=\mu_2=\mu_3=\mu$, then by differentiating the equations (2) with respect to $t$, and substituting for $d\alpha/dt$, etc., from the equations (1), we obtain the equations

$$\left.\begin{aligned} \mu K_1\frac{d^2P}{dt^2}&=\nabla^2P-\frac{d\Delta}{dx}\\ \mu K_2\frac{d^2Q}{dt^2}&=\nabla^2Q-\frac{d\Delta}{dy}\\ \mu K_3\frac{d^2R}{dt^2}&=\nabla^2R-\frac{d\Delta}{dz}\end{aligned}\right\}\qquad (3),$$

where

$$\Delta=\frac{dP}{dx}+\frac{dQ}{dy}+\frac{dR}{dz}.$$

In order to determine how the velocity of propagation of a plane wave depends on the direction of propagation in the medium, let us take the case of a plane wave travelling in the direction $l$, $m$, $n$, where $l$, $m$, $n$ are the direction cosines of the wave normal, then in general the electric and magnetic forces at any point, $x$, $y$, $z$ may be expressed as periodic functions of $lx + my + nz - vt$ where $v$ is the wave velocity. Hence if we take a simple harmonic disturbance we may write

$$P = P_0 \sin(lx + my + nz - vt),$$

with similar expressions for Q and R. From this it follows at once that

$$\frac{d^2P}{dx^2} = -l^2P, \quad \frac{d^2P}{dt^2} = -v^2P, \quad \frac{d^2Q}{dxdy} = -lmQ, \text{ etc.}$$

Consequently, if we write $\mu K_1 = 1/v_1^2$, $\mu K_2 = 1/v_2^2$, $\mu K_3 = 1/v_3^2$, where $v_1$, $v_2$, $v_3$ are the three principal velocities, then the equations (3) become

$$\left.\begin{aligned}(v^2 - v_1^2)P + lv_1^2(lP + mQ + nR) = 0\\ (v^2 - v_2^2)Q + mv_2^2(lP + mQ + nR) = 0\\ (v^2 - v_3^2)R + nv_3^2(lP + mQ + nR) = 0\end{aligned}\right\} \quad (4).$$

Dividing these equations by $v^2 - v_1^2$, $v^2 - v_2^2$, $v^2 - v_3^2$, and adding them together, after multiplying by $l$, $m$, $n$ respectively, we have

$$1 + \frac{l^2v_1^2}{v^2 - v_1^2} + \frac{m^2v_2^2}{v^2 - v_2^2} + \frac{n^2v_3^2}{v^2 - v_3^2} = 0,$$

and, remembering that $1 = l^2 + m^2 + n^2$, this may be written in the form

$$\frac{l^2}{v^2 - v_1^2} + \frac{m^2}{v^2 - v_2^2} + \frac{n^2}{v^2 - v_3^2} = 0 \quad (5).$$

Thus the velocity of propagation of an electromagnetic wave in an electrically crystalline medium is determined by the same equation as that deduced by Fresnel for the propagation of a light wave in a crystal (Art. 193), and, starting from this stage, all the results obtained in chap. xii. for doubly refracting crystals may be expressed in terms of the electromagnetic theory.

Now P, Q, R are proportional to the direction cosines of the electric force, and in the foregoing equations we see that $lP + mQ + nR$ cannot be zero unless $v = v_1 = v_2 = v_3$—that is, unless the medium is isotropic. But if $lP + mQ + nR$ is not zero the electric force is not at right angles to the wave normal, and consequently not in the wave front. It is easily seen, however, that the electric displacement is in the wave front. For from equations (4) it is obvious that

$$\frac{P}{v_1^2} : \frac{Q}{v_2^2} : \frac{R}{v_3^2} = \frac{l}{v^2 - v_1^2} : \frac{m}{v^2 - v_2^2} : \frac{n}{v^2 - v_3^2}.$$

But $P/v_1^2 : Q/v_2^2 : R/v_3^2 = K_1P : K_2Q : K_3R$, and these latter are the components of the electric displacement, and are therefore proportional to the direction cosines of the displacement. Denoting these by $l'$, $m'$, $n'$, we have

$$l' : m' : n' = \frac{l}{v^2 - v_1^2} : \frac{m}{v^2 - v_2^2} : \frac{n}{v^2 - v_3^2} \quad (6).$$

Consequently by equation (5) we have

$$ll' + mm' + nn' = 0,$$

that is, the electric displacement is in the wave front or at right angles to the wave normal. This relation may also be deduced very easily from the equations (2), assuming $\alpha$, $\beta$, $\gamma$ to be functions of $lx + my + nz - vt$.

Comparing the relations (6) with those of Art. 195 we see that the direction of the electric displacement is the same as that of the vibration considered in Fresnel's theory—that is, the electric displacement is perpendicular to the plane of polarisation, while the magnetic displacement lies in that plane, and is therefore the vibration considered by MacCullagh.

On the other hand, if we take K to be the same in all directions while $\mu$ is variable, we obtain differential equations for $\alpha$, $\beta$, $\gamma$, which are exactly the same as those from which we have deduced the preceding results for P, Q, R. We conclude, therefore, that in a medium which is magnetically æolotropic, but electrically isotropic, the magnetic displacement is in the wave front and at right angles to the plane of polarisation, while the electric displacement is in the plane of polarisation, as in MacCullagh's theory.

Finally, when the medium is neither electrically nor magnetically isotropic the equation of the wave surface becomes

$$\left(\frac{x^2}{K_1}+\frac{y^2}{K_2}+\frac{z^2}{K_3}\right)\left(\frac{x^2}{\mu_1}+\frac{y^2}{\mu_2}+\frac{z^2}{\mu_3}\right)$$
$$-\frac{x^2}{\mu_1 K_1}\left(\frac{1}{\mu_2 K_3}+\frac{1}{\mu_3 K_2}\right)-\frac{y^2}{\mu_2 K_2}\left(\frac{1}{\mu_3 K_1}+\frac{1}{\mu_1 K_3}\right)-\frac{z^2}{\mu_3 K_3}\left(\frac{1}{\mu_1 K_2}+\frac{1}{\mu_2 K_1}\right)+\frac{1}{\mu_1\mu_2\mu_3 K_1 K_2 K_3}=0.$$

This is the surface enveloped by the wave planes. It is of the fourth degree, and each of the co-ordinate planes intersects it in a pair of ellipses. In one of these planes the ellipses intersect each other in four conical points. In all ordinary crystals $\mu_1 = \mu_2 = \mu_3 = 1$ very approximately, and the foregoing general equation reduces to Fresnel's well-known form.

## APPENDIX I.

The following tables are taken from Watts's *Dictionary of Chemistry*. The indices are for the yellow rays, except those of Wollaston, which are for the extreme red :—

INDICES OF REFRACTION OF SOLIDS

| Solid. | Index. | Observer. |
|---|---|---|
| Lead Chromate . . . . | 2·5 to 2·97 | Brewster. |
| Diamond . . . . . | 2·47 to 2·75 | Br. and Rochon. |
| Phosphorus . . . . . | 2·224 | Brewster. |
| Glass of Antimony . . | 2·216 | ,, |
| Sulphur (native) . . . | 2·115 | ,, |
| Zircon . . . . . . | 1·95 | Wollaston. |
| Lead Nitrate . . . . | 1·866 | Herschel. |
| Lead Carbonate . . . | 1·81 to 2·08 | Brewster. |
| Ruby . . . . . . | 1·779 | ,, |
| Felspar. . . . . . | 1·764 | ,, |
| Tourmaline . . . . . | 1·668 | ,, |
| Topaz (colourless) . . . | 1·610 | Biot. |
| Beryl . . . . . . | 1·598 | Brewster. |
| Tortoise-shell . . . . | 1·591 | ,, |
| Emerald . . . . . | 1·585 | ,, |
| Flint Glass . . . . . | 1·57 to 1·58 | Br. and Woll. |
| Rock-crystal (least) . . | 1·547 | Wollaston. |
| Rock-salt . . . . . | 1·545 | Newton. |
| Apophyllite . . . . . | 1·543 | Brewster. |
| Colophony . . . . . | 1·543 | Wollaston. |
| Sugar . . . . . . | 1·535 | ,, |
| Phosphoric Acid . . . . | 1·534 | Brewster. |
| Copper Sulphate . . . . | 1·531 to 1·552 | , |
| Canada Balsam . . . . | 1·532 | Young. |
| Citric Acid . . . . . | 1·527 | Brewster. |
| Crown Glass . . . . . | 1·525 to 1·534 | ,, |
| Nitre . . . . . . | 1·514 | ,, |
| Plate Glass . . . . . | 1·514 to 1·542 | ,, |
| Spermaceti . . . . . | 1·503 | Young. |
| Crown Glass . . . . . | 1·5 | Wollaston. |
| Potassium Sulphate . . | 1·5 | Brewster. |
| Ferrous Sulphate . . . . | 1·494 | ,, |
| Tallow ; Wax . . . . | 1·492 | Young. |
| Magnesium Sulphate . . | 1·488 | Brewster. |
| Iceland Spar (greatest) . . | 1·654 | Malus. |
| Obsidian . . . . . | 1·488 | Brewster. |
| Gum . . . . . . | 1·476 | Newton. |
| Borax . . . . . . | 1·475 | Brewster. |
| Alum . . . . . . | 1·457 | Wollaston. |
| Fluorspar . . . . . | 1·436 | Brewster. |
| Ice . . . . . . | 1·310 | Wollaston. |
| Tabasheer . . . . . | 1·1115 | Brewster. |

# APPENDIX II.

## Indices of Refraction of Liquids

| Liquid. | Index. | Observer. |
|---|---|---|
| Sulphide of Carbon | 1·678 | Brewster. |
| Oil of Cassia | 1·631 | Young. |
| Bitter Almond Oil | 1·603 | Brewster. |
| Nut Oil | 1·500 | ,, |
| Linseed Oil | 1·485 | Wollaston. |
| Oil of Naphtha | 1·475 | Young. |
| Rape Oil | 1·475 | ,, Br. |
| Olive Oil | 1·470 | Brewster. |
| Oil of Turpentine | 1·470 | Wollaston. |
| Oil of Almonds | 1·469 | ,, |
| Oil of Lavender | 1·457 | Brewster. |
| Sulphuric Acid (sp. gr. 1·7) | 1·429 | Newton. |
| Nitric Acid (sp. gr. 1·48) | 1·410 | Young, Woll. |
| Solution of Potash (sp. gr. 1·41) | 1·405 | Fraunhofer. |
| Hydrochloric Acid (concentrated) | 1·410 | Biot. |
| Sea Salt (saturated) | 1·375 | ,, |
| Alcohol (rectified) | 1·372 | Herschel. |
| Ether | 1·358 | Wollaston. |
| Alum (saturated) | 1·356 | Herschel. |
| Human Blood | 1·354 | Young. |
| White of Egg | 1·351 | Euler (jun.) |
| Vinegar (distilled) | 1·372 | Herschel. |
| Saliva | 1·339 | Young. |
| Water | 1·336 | Woll., Br. |

# INDEX

*The Numbers refer to the Pages.*

THE END

*Printed by* R. & R. CLARK, LIMITED, *Edinburgh.*