Newton and the birth of the theory of colours

by Claude Guthmann, emeritus professor of physics
at Paris-Denis Diderot University (Paris VII)

***Figure 1*: German stamp (1993) and Isaac Newton's signature.** *On the stamp we can see the law of dynamics* **F** = mγ *(expressed in the form* mΔv = FΔt*), as well as the spectral decomposition of light, which the present article examines.*

This letter of 6 February 1672 is the first publication of Newton's discoveries concerning the decomposition of white light by a prism and the theory of colours he developed on this basis.

> *[H]aving darkened my chamber, and made a small hole in my window-shuts …*[1]

Thus Newton sets up his experiment and his theory of light and colours. He then places a glass prism against the hole and observes the refraction of the white light on the opposite wall. He sees the dispersed colours and is surprised to find that they form an oblong; he was, he says, expecting them to be circular, like the shape of the hole. In fact, what he saw was the sum of an infinite number of circular images of the sun, corresponding to the various "colours" that

1. --Trans. The citations from Newton's letter are taken from the Newton Project web page: http://www.newtonproject.ox.ac.uk/view/texts/normalized/NATP00006

we know to be refracted differently by the prism in the direction perpendicular to its edge. If the colours were all refracted in the same way, in a circular image of the hole, there would be no separation of colours! Newton also observed a coloured spectrum that was five times as long as it was wide – an oblong or elliptical form rather than a circular one.

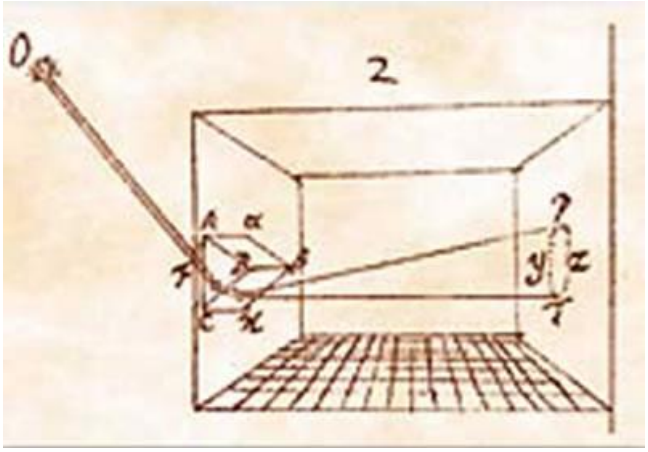
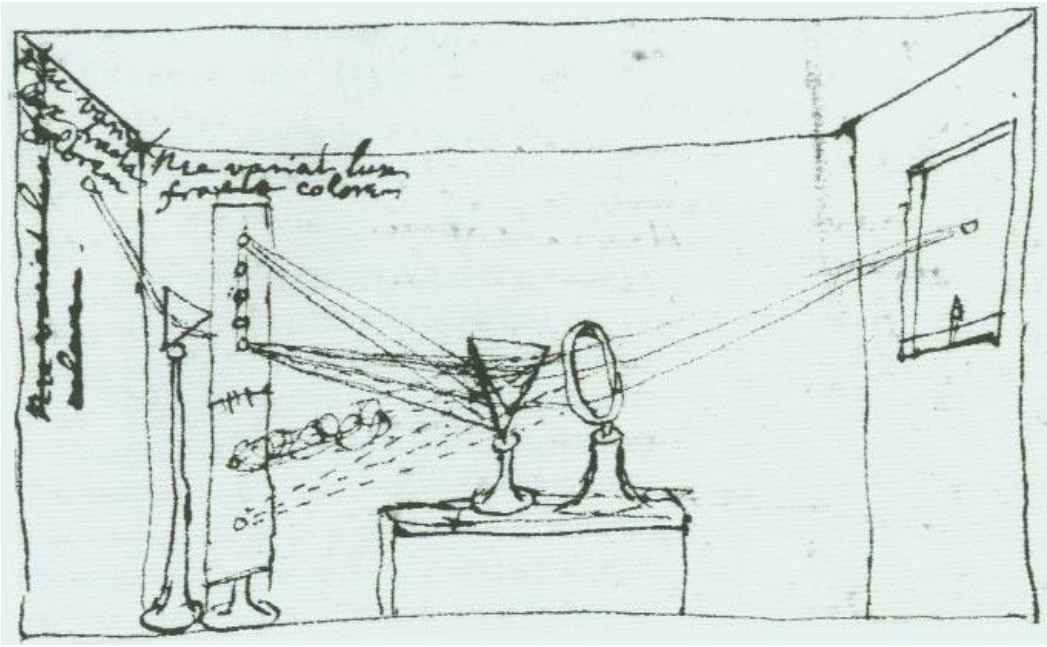


Figure 2 (above) and 2a (below): Drawings attributed to Newton.

Fig. 2: *This is the diagram of the experiment: the prism is placed in front of the window, and the oblong mark is observed on the wall to the right.*

Fig. 2a: *The set-up here is slightly different from that described in the letter: the prism is on a table instead of being placed in front of the hole in the window shutter, and the image of the latter is reflected by a lens onto the prism. The oblong figure can be seen on the left-hand screen, between the two outermost rays. The circular shape, corresponding to the angular diameter of the sun (approximately ½°, i.e. 30′), transforms into this shape, which is five times as long (2°49′, i.e. almost 3°) (image: Warden and Fellows, New College, Oxford)*



Newton now begins to seek the origin of this phenomenon. He rules out effects due to the geometry of the hole or variation in the thickness of the glass:

> *I could scarce think, that the various Thickness of the glass, or the termination with shadow or darkness, could have any Influence on light to produce such an effect; yet I thought it not amiss, first to examine those circumstances, and so tryed, what would happen by transmitting light through parts of the glass of divers thicknesses, or through holes in the window of divers bignesses, or by setting the Prisme without so, that the light might pass through it, and be refracted before it was terminated by the hole: But I found none of those circumstances material. The fashion of the colours was in all these cases the same.*

His next approach, to test *whether by any unevenness in the glass, or other contingent irregularity, these colours might be thus dilated*, was to take a second prism identical to the first and place it contiguous to the other, so as to obtain the opposite effect. And, in effect, the exiting beam of light was now identical to the incident beam, rather than dispersed:

> *the light, which by the first Prisme was diffused into an* oblong *form, was by the second reduced into an* orbicular *one with as much regularity, as when it did not at all pass through them.*

From this he concludes that:

> *what ever was the cause of that length, 'twas not any contingent irregularity.*

This signifies in particular that the separation of colours effect does not occur in one given prism but not in another. Today we would say that the effect is reproducible.

Newton now turned to a more quantitative experimental study. By measuring the dimensions of the image obtained, he found that the angle subtended by the dilated image of the colours equals 2° 49′, whereas the angle that subtends its breadth and corresponds to the apparent diameter of the sun equals 31′. He then calculated the refraction of two rays proceeding from diametrically opposed points of the sun's discus and found that the emergent rays must indeed subtend 31′. In effect, he took equal angles of incidence and emergence to and from the prism (54°4′); this is the minimum deviation possible from the prism and the rays with different angular incidence will leave the prism with almost the same angular difference.

Having made these observations, I first computed from them the refractive power of that glass, and found it measured by the ratio of the sines, 20 to 31.

For this he used the law of sines: $n_1 \sin i_1 = n_2 \sin i_2$, where i_1 and i_2 are the angles of incidence and refraction with respect to the surface of separation of two index media n_1 and n_2.

The minimum of deviation of a prism

Newton uses a prism with an apex angle (*A* below) equal to 63° and sets the experiment at the minimum of deviation, that is to say the angle of incidence *i* that makes the deviation *D* minimal.

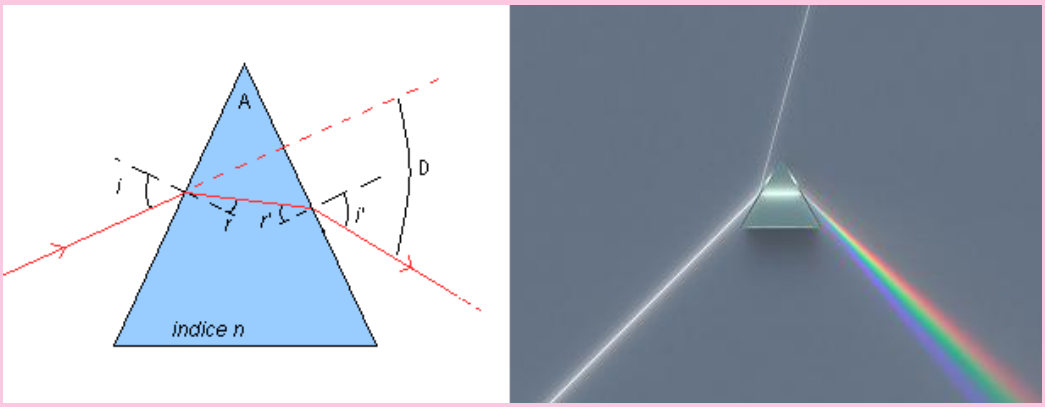


***Figure 3*: Refraction of light through a prism.** *On the left, we can see a significant deviation of the initial ray, i.e. over 90° (image: Wikimedia Commons).*

We'll skip the equations of deviation *D* as a function of incidence *i*, as they are quite complex; they are based on the three formulas $A = r + r'$ (the very simple geometry of the triangle), and $\sin i = n \sin r$, $\sin i' = n \sin r'$ (Descartes' laws), with *n* being the index of refraction of the glass. There exists a "minimum of deviation" for a certain angle of incidence *i*, dependent on the angle of prism *A*. In this case, we see that the situation is symmetrical in *i* and *i'* – if the light were to arrive in the other direction it would also be a minimum; we therefore have $i = i'$.
The minimum of deviation thus occurs for an angle of incidence such as:

$$\sin i = n \sin(A/2)$$

Newton notes that his prism has an angle of $A = 63°12'$. With the index of refraction that he calculates, $n = 1.55$, the calculation gives the minimal incidence of deviation as 54°12' (scarcely different from the value given by Newton, 54°24').[2]

Newton thus calculates the index as 1.55, an entirely reasonable value for glass. He then immediately recalculates the deviations of two incident rays

2. A simulation of the minimum of deviation can be found here, on the website of the University of Nantes.

differing by 31′ and again finds an angle of 31′ between the emergent rays when they exit the prism. From this he concludes that such a calculation does not give the – distinctly greater – 2°49′ observed.

He goes back to his prism and rotates it a little on its axis in both directions, so as to vary the angle of incidence of the light by 4 to 5 degrees. He observes:

> *the Colours were not thereby sensibly translated from their place on the wall, and consequently by that variation of Incidence, the quantity of Refraction was not sensibly varied.*

The difference in the incidence of the rays originating from various parts of the sun cannot explain the angle of 2°49′ observed.

Another reason was needed. Newton thus considers whether the rays might follow a curved trajectory after their trajectory through the prism, and whether, depending on their greater or lesser curvature, they might reach different points of the wall. He remembers having seen a tennis ball hit sidelong that followed just such a trajectory. This is in line with his hypothesis about the corpuscular nature of light: if light rays were globular bodies and were able to acquire a circular movement by passing from one medium to another at a non-normal incidence at the interface, then could thus undergo a greater resistance from the ether surrounding them on the side where the movements accumulate, and thus be deviated towards the other side. Concerning the light corpuscles, in order to replace the air that the tennis ball might rub against Newton has recourse to a matter dear to physicists of his age: ether. An independent medium for the propagation of vibrations, ether, which was invisible but nonetheless present, played a preponderant role in vibratory theories of the 17th and 18th century.

But, despite the plausibility of his hypothesis, Newton quickly takes on board the evidence: he observes no curvature of the rays. He now carries out what he calls the *Experimentum Crucis*.

He arranges two screens, one very close behind the window prism, and bearing a small hole through which light can travel. The light reaches the second screen, which is also pierced by a small hole, so that some of the incident light traverses it. He now places a second prism behind the second screen so that the light, passing through the two screens, can also traverse the prism and again be refracted before reaching the wall.

The drawing below depicts the experiment in schematic form:

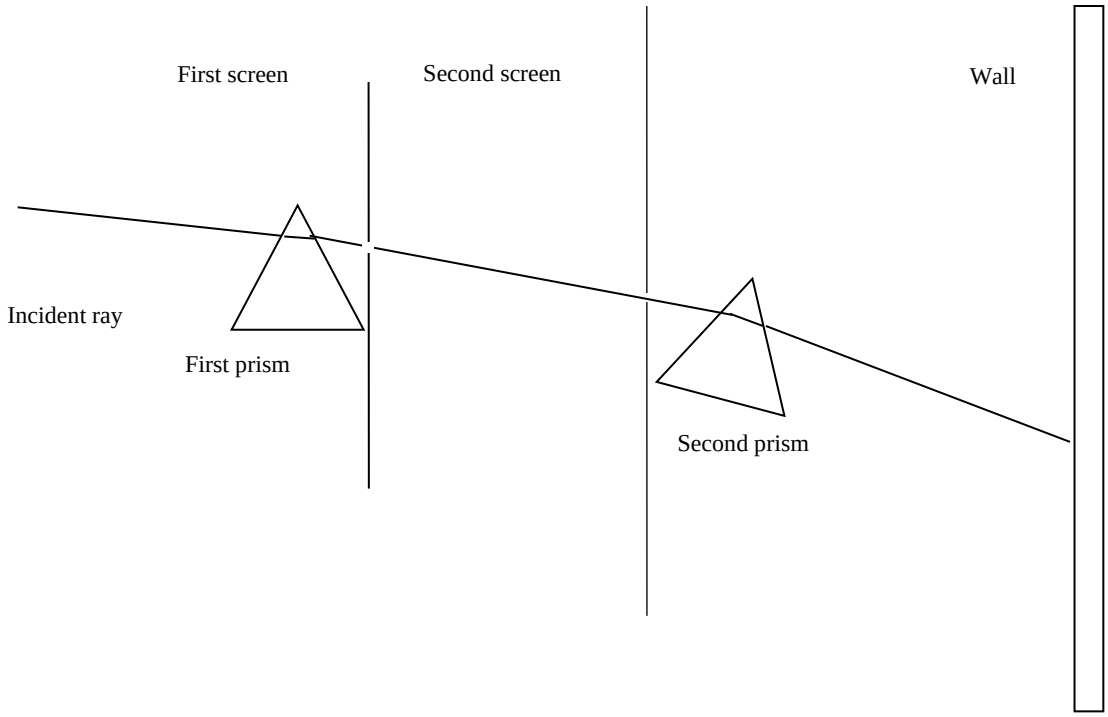
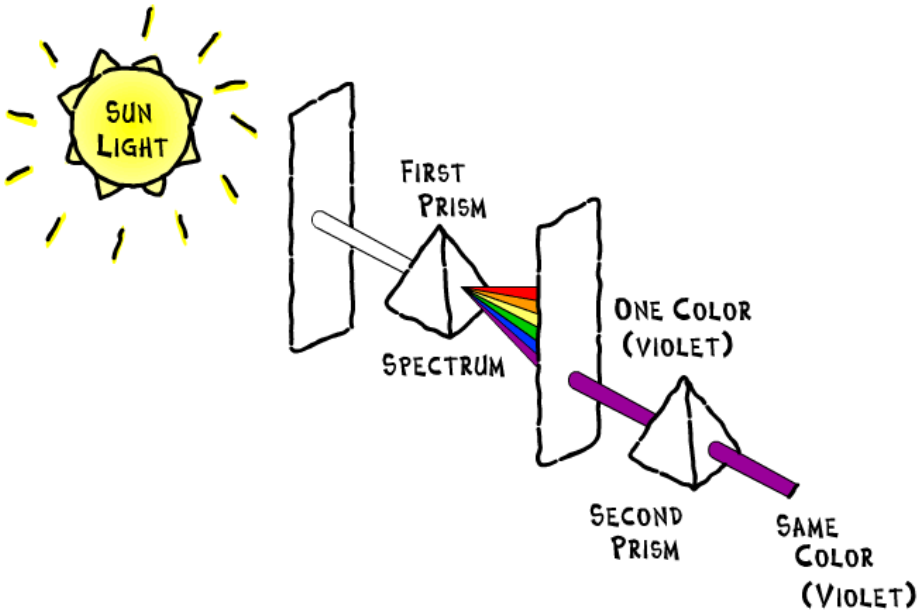


***Figure 4*: The "Experimentum crucis".** *Above, diagram of Newton's experiment (author's illustration). Below, the stylised result, in colour (Michigan State University). The left-hand side of Figure 2a above (Newton's drawing) also shows the* Experimentum Crucis*: the screen is pierced with a single hole, which allows a colour to pass through. This colour then passes through the second prism, though the white light is not reconstituted.*



Newton now rotates the first prism:

to make the several parts of the Image, cast on the second board, successively pass through the hole in it, that I might observe to what places on the wall the second Prisme would refract them.

In other words, he uses the second hole to select some of the colours separated by the first prism. And, he states, he has seen *that the light*

> *tending to that end of the Image, towards which the refraction of the first Prisme was made, did in the second Prisme suffer a Refraction considerably greater then the light tending to the other end,*

which means that purple is deviated more than the red in the second as well as the first prism. Newton concludes that the true cause of the elongation of the image is as follows: the light consisted in rays of differing refrangibility, which, independently of their difference in incidence, were transmitted towards different parts of the wall, depending on their degree of refringence. He is perfectly correct, as we now know, but the experiment is not as crucial as he says. In fact, the experiment had merely shown that all equally refracted rays invariably have the same colour and that differently refracted rays are different colours or, to summarise, prisms always deviate purple more than red. Newton's brilliant intuition was to recognise that this is not an intrinsic property of prisms but that white light is composed of radiations of different "colours". Today we say that white light is the resultant of radiations of different frequencies and that the prism produces a spectral decomposition of these. Fundamentally, Newton puts forward a model that is well aligned with the experiment. Over the years, this model would establish itself given the number of physical consequences that it entailed, and which were always corroborated.

As for the corpuscular model of light, which is different from Huygens's undulatory model that he mentions earlier, there is agreement only as concerns the reflection and refraction of light. It would be altered and confirmed only in 1905 in an article by Einstein that interprets the photoelectric effect and presents the undulatory and corpuscular character of light quanta, later called photons, for the first time.

CHROMATIC ABERRATION

Let's return to the immediate conclusions Newton draws from his discovery. He writes:

> *When I understood this, I left off my aforesaid Glass works; for I saw, that the perfection of Telescopes was hitherto limited, not so much for want of glasses truly figured according to the prescriptions of Optick Authors, (which all men have hitherto imagined,) as because that Light it self is a* Heterogeneous mixture of differently refrangible Rays. *So that,*

were a glass so exactly figured, as to collect any one sort of rays into one point, it could not collect those also into the same point, which having the same Incidence upon the same Medium are apt to suffer a different refraction.

This is an excellent definition of what we now call chromatic aberration. And although, following this remark, scientists learnt shortly after Newton's death how to more or less correct it using optical instruments known as achromats (lenses made of glasses with different degrees of dispersion, affixed together), this was not yet the case in 1672. Newton immediately makes a suggestion to remedy this: producing optic instruments by reflection. Indeed, *the Angle of Reflection of all sorts of Rays was equal to their Angle of Incidence*, he says. To bring

Optick instruments … to any degree of perfection imaginable

it was necessary for

a Reflecting *substance* [to] *be found, which would polish as finely as Glass, and* reflect *as much light, as glass* transmits*, and* [that] *the art of communicating to it a* Parabolick *figure be also attained.*

This last requirement is necessary to ensure the stigmatism of the mirror for rays coming from infinity.

Two years later, thanks to a simple polishing method suitable for metals, he produced a telescope with a metallic mirror:

I began to try, what might be effected in this kind, and by degrees so far perfected an Instrument (in the essential parts of it like that I sent to London*,) by which I could discern Jupiters 4 Concomitants, and shewed them divers times to two others of my acquaintance.*

He also envisaged making an optical reflection microscope, again to get around the issue of chromatic aberration.

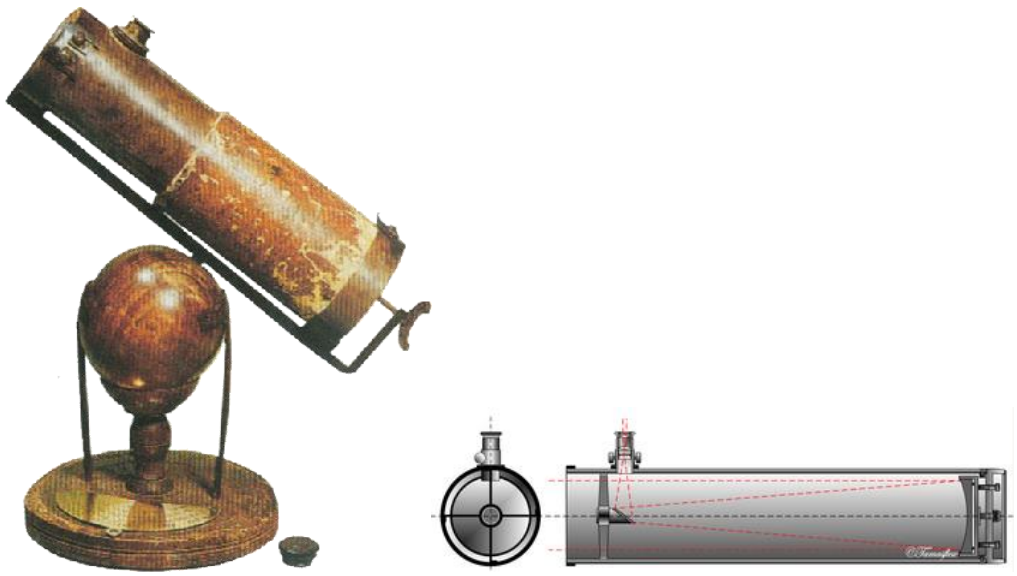


***Figure 5*: Newton's telescope, 1672 (left) and diagram of the light rays inside (right).** *On the left we can see the eyepiece, which is perpendicular to the optical tube. The telescope works by reflection onto a parabolic mirror (back right), unlike astronomical telescopes, in which the rays are refracted onto lenses (right-hand drawing: Wikimedia Commons, author: Szőcs Tamás).*

A THEORY OF COLOURS IN THIRTEEN POINTS

After this digression on the chromatic aberration of optical instruments, Newton returns to his subject. He now takes it as a given that

> *Light is not similar, or homogeneal, but consists of* difform *Rays, some of which are more refrangible than others.*

He now applies this to a 13-point theory of colours, followed by experimental examples.

1°) Light rays differ not only on account of their degree of refrangibility, but also on account of their colour.

> *Colours are not* Qualifications of Light*, derived from Refractions, or Reflections of natural Bodies (as 'tis generally believed).*

Moreover, not only are there special kinds of rays for the most important colours, but on the contrary rays for all the intermediary stages. Today we would say: the refraction of light by a given substance depends on the wavelength.

2°)

> *To the same degree of Refrangibility ever belongs the same colour, and to the same colour ever belongs the same degree of Refrangibility. The* least Refrangible *Rays are all disposed to exhibit a* Red *colour, and contrarily those Rays, which are disposed to exhibit a* Red *colour, are all the least*

refrangible: So the most refrangible *Rays are all disposed to exhibit a deep* Violet Colour*, and contrarily those which are apt to exhibit such a violet colour, are all the most Refrangible. And so to all the intermediate colours in a continued series belong intermediate degrees of refrangibility.*

3°)

The species of colour, and degree of Refrangibility proper to any particular sort of Rays, is not mutable by Refraction, nor by Reflection from natural bodies, nor by any other cause, that I could yet observe. When any one sort of Rays hath been well parted from those of other kinds, it hath afterwards obstinately retained its colour, notwithstanding my utmost endeavours to change it. I have refracted it with Prismes, and reflected it with Bodies, which in Day-light were of other colours; I have intercepted it with the coloured film of Air interceding two compressed plates of glass; transmitted it through coloured Mediums, and through Mediums irradiated with other sorts of Rays, and diversly terminated it; and yet could never produce any new colour out of it. It would by contracting or dilating become more brisk, or faint, and by the loss of many Rays, in some cases very obscure and dark; but I could never see it changed in specie*.*

Newton insists on the intrinsic character of the property of colour, which we would today term the wavelength or frequency of light. To back up what he is saying, he cites numerous examples.

4°) It is possible to obtain different colours by mixing given colour rays. By adding together differently coloured rays, one obtains an intermediary colour. This in turn may be decomposed by refraction, for example, thereby retrieving the original colour rays of which it is composed. Newton cites the analogy of blue and yellow powders which, once mixed, appear green to the naked eye and yet viewed through a microscope are juxtaposed but remain blue and yellow.

5°)

There are therefore two sorts of Colours. The one original and simple, the other compounded of these. The Original or primary colours are, Red, Yellow, Green, Blew*, and a* Violet-purple*, together with Orange, Indico, and an indefinite variety of Intermediate gradations.*

And there we have the foundation of the theory of colours.

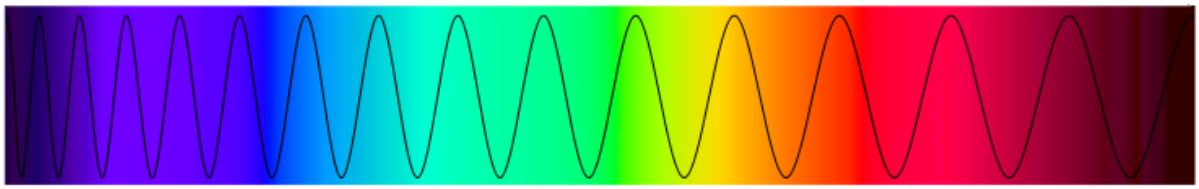


***Figure 6*: The spectrum of visible light.** *Newton's decomposition into seven original colours is arbitrary, more or less corresponding to the colours visible to the naked eye. We now know that the spectrum of light is continuous, corresponding to the continuity of electromagnetic frequencies from 400 nm (purple) to 700 nm (red).*

6°) This paragraph completes the previous one. It states that, if one arranges two proximate colours in the series of colours separated by the prism, we obtain the colour appearing in the preceding series between these two colours.

7°) The recomposition of all the primary colours produces white light. By making them converge and mixing them together again, as in the original light, we obtain white light, a fact that Newton cannot help finding *surprising and wonderful*. He also remarks that the white light thereby obtained is no different from sunlight, unless the glass used in the convergence optical instrument is slightly tinted, which gives a slight hue to the recomposed light.

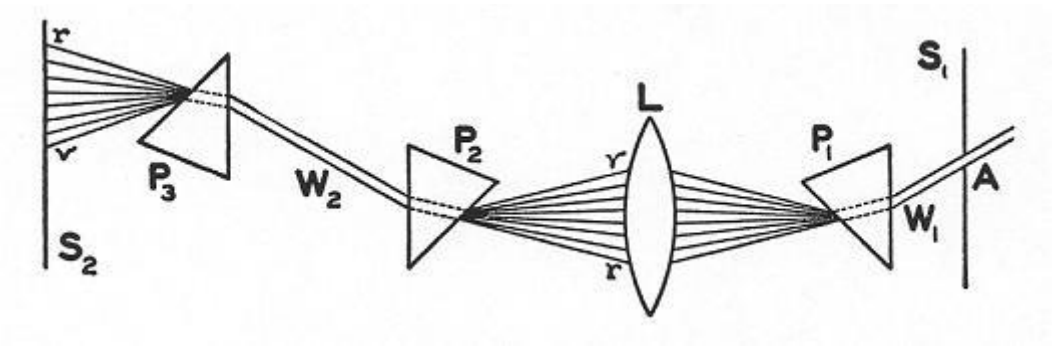


***Figure 7*: Recomposition of light.** *The light arriving from the right (W_1) is decomposed by prism P_1. The lens L refocuses it on prism P_2, which is positioned symmetrically to P_1. Prism P_2 recomposes the coloured spectrum into a white light W_2 analogous to W_1. This beam can once again be decomposed onto prism P_3, thereby displaying the colours on the screen, from purple (the most refracted, below) to red (the least refracted, above) (image: Appalachian State University).*

8°)

> *Hence therefore it comes to pass, that* Whiteness *is the usual colour of* Light*; for, Light is a confused aggregate of Rays indued with all sorts of Colors, as they are promiscuously darted from the various parts of luminous bodies.*

This sentence, which might sound a bit naïve to us, could today be stated as: white light has a spectrum ranging from purple to red.

9°) On the basis of these statements, Newton offers an interpretation of the separation of the colours by the prism:

> *For, of the Rays, constituting the incident light, since those which differ in Colour proportionally differ in Refrangibility,* they *by their unequall refractions must be severed and dispersed into an oblong form in an orderly succession from the least refracted Scarlet to the most refracted Violet.*

Today we would say that the index of the prism glass is a function of wavelength, i.e. the colour. For each wavelength there is a corresponding and different deviation by the prism.

10°) He now gives an explanation of the rainbow. We now know that

a) the light refracted when it penetrates the raindrop is partially reflected at the back of the raindrop (point B in the diagram) and is again refracted when it leaves (point C). The resulting deviation is stronger for purple than for red. This is the origin of the primary – and most intense – arc, which occupies an angular band of 40° to 42° in the opposite direction to the sun.

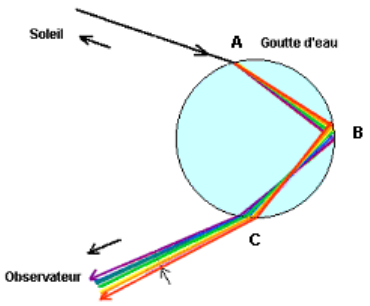


***Figure 8*: The raindrop acts like a prism.** *It should be noted that in B the rays refracted by the raindrop (not represented here) also leave: this is the opposite phenomenon to A – the white light recomposes into a single beam. On the other hand, the exit from the raindrop in C will be different from the exit from the raindrop in B: after the first reflection in B, the refraction in C will be different depending on the colours, and the single light beam is not reconstituted.*

b) A double reflection of the sunlight may occur within the raindrops. A secondary arc thus appears at an angle of 50° to 53°. Due to the additional reflection, the arc is less bright and the order of the colours is inversed compared to the primary arc, with the blue on the outside being less deviated than the red within.

> *Why the Colours of the* Rainbow *appear in falling drops of Rain, is also from hence evident. For, those drops, which refract the Rays, disposed*

to appear purple, in greatest quantity to the Spectators eye, refract the Rays of other sorts so much less, as to make them pass beside it; and such are the drops on the inside of the Primary *Bow, and on the outside of the* Secondary *or Exteriour one. So those drops, which refract in greatest plenty the Rays, apt to appear red, toward the Spectators eye, refract those of other sorts so much more, as to make them pass beside it; and such are the drops on the exteriour part of the* Primary*, and interiour part of the* Secondary *Bow.*

This statement, although essentially correct, is a little complicated and suggests that raindrops do not all refract and reflect light in the same way, which is untrue.

11°) When substances reflect a wavelength (*a sort of light*, Newton calls it) and transmit another, their colour will vary according to their position vis-à-vis the light source and the observer. Thus an infusion made from the wood *Lignum nephreticum* has a changing opal appearance, depending on the angle at which it is lit. The same is true of gold leaf and fragments of coloured glass.

12°) An experiment due to Hooke consisted in sticking together two wedge-shaped glass vessels, one of which contained a red liquid and the other a blue liquid, and observing that together they no longer transmitted light. Newton interprets this perfectly by saying that

if one transmitted only red, and the other only blew, no rays could pass through both.

The red wedge does indeed absorb all radiation that is not red, and nor will the latter be transmitted by the blue wedge, which absorbs all radiation that is not blue.

13°)

I might add more instances of this nature, but I shall conclude with this general one, that the Colours of all natural Bodies have no other origin than this, that they are variously qualified to reflect one sort of light in greater plenty then another.

The entire theory of colours lies in this statement. And to corroborate what he is saying, Newton adduces the result of his experiments in the dark room. Illuminating various bodies with lights of different colours, he saw that all the bodies were always illuminated in the colour of the light shining on them:

but yet with this difference, that they are most brisk and vivid in the light of their own day-light-colour.

He then cites the examples of minium and azurite, which are situated at each extreme of the colour spectrum.

***Figure 9*: On the left, azurite ore; on the right, minium powder.** *Azurite (blue in colour) is a copper carbonate that has been known since Antiquity – traces of it have been found in Egyptian ointment pots. Minium is a lead tetroxide with the formula Pb_3O_4, and has also long been used as a wall-painting pigment.*

And so,

> *that this is the intire and adequate cause of their colours, is manifest, because they have no power to change or alter the colours of any sort of Rays incident apart, but put on all colours indifferently, with which they are inlightned.*

Newton then argues that colours are *qualities*, that is to say properties of light, in order to infer that light is indeed a substance. This is in line with a corpuscular conception of light. *But he immediately adds that determining more absolutely what light is, the way in which it refracts and the way in which it produces the sensation of colours in our minds is far from straightforward*.

@@@@@@@

He now comes to a description of the final experiment. He is still using a darkened room into which an orifice pierced through a window shutter allows sunlight to pass. A prism is placed against the orifice through which the light enters to form a coloured oblong image on the opposite wall of the room. In the trajectory of the differently coloured rays one places a convergent lens with an opening (15.2 cm) large enough to capture all the refracted rays and make them converge on its focal spot, situated around a quarter of a metre further on. On a screen made out of a sheet of white paper placed at this point, one observes a

mark of white light resulting from the mix of the refracted colours. Moving this screen around the focal point, one will see

> *to what distance the whiteness is most perfect, but also see, how the colours gradually convene, and vanish into whiteness, and afterwards having crossed one another in that place where they compound Whiteness, are again dissipated, and severed, and in an inverted order retain the same colours, which they had before they entered the composition.*

We could consider this experiment to be "crucial" too, in Newton's sense of the term, but he makes no mention of it. He describes all the parameters of the experiment with great precision and stresses the need for an extremely dark room to avoid any parasitic diffused light that might distort the separation of colours.

CONCLUSION

This brings us to the end of this surprising letter. Newton has introduced the equivalence between the colour of the rays and their various degrees of refrangibility in a given index medium. He has given a coherent theory of the colours of natural bodies, due to a selective reflection that he has demonstrated by successively projecting the various single colours onto the bodies being observed in a darkened room. Moreover, his discovery would provide the foundation of spectroscopy, a central branch of modern physics. Lastly, he is not afraid to apply it to a practical problem: the limited resolution of optical apparatuses by what we now call chromatic aberration. And, to free himself from this, he has designed and constructed the first reflection-type telescope, which now bears his name.

To conclude, let's note the extent to which this letter, in which Newton draws together hypothesis and experiments with extraordinary precision and clarity, is a perfect illustration of the experimental method. And the fact that he had no trouble refuting the critiques which his theory aroused at the time, particularly on the part of Huygens.[3]

■

(October 2010)

(Translated by Helen Tomlinson, published July 2017)

3. L. Rosenfeld, *Isis*, Vol. 9, No. 1 (Feb. 1927), pp. 44–65.