Chapter 1

Understanding the Nature of Light

The nature of light has fascinated philosophers, as well as ordinary people, for thousands of years. The light seemed something ordinary and at the same time something very mysterious, magical. According to today's criteria, a truly scientific approach to the nature of light was initiated about 350 years ago by two great scholars - the Englishman Isaac Newton and the Dutchman Christian Huygens. By then, three firmly established facts about light were already known:

- that light propagates along straight lines – the evidence of that are shadows;

- that if light is reflected from flat mirror or any other flat surface (e.g., water surface), the angle between the surface and the incident light rays, and the angle between the surface and the reflected light rays – are exactly equal: an evidence for that, for instance, is that the mirror image of an object is of t he same size as the object;

- that light, when passing through the interface between two transparent media, is refracted;

- that the edges of a shadow are not perfectly "sharp" – as if some part of light passing near the edges of an shadow-forming object is slightly "deflected". It is best seen when light passes through a narrow slit – the image of the slit on a screen is wider than the slit.

Both Newton and Huygens presented explanations for the above phenomena – but completely different explanations!

One of the greatest achievements of Newton was the formulation of the laws of mechanics – which tell us how **material objects**, either tiny ones, or huge ones, behave when forces are acting on them – and, in particular, when there is no such force. Furthermore, he discovered the law of universal gravity – saying that between any two material objects there is always an attractive force, proportional to the objects' masses, and inversely proportional to the square of the distance between them.

Newton realized that if light were a stream if tiny material particles, than the laws of mechanics he had discovered would be able to explain all the four phenomena listed above! Let us follow his reasoning, for each of them.

The Newton's *First Law of Dynamics* states that if there is no force acting on a material object, it either is at rest, or moves with a *constant velocity along a straight line.* So, if light is a stream of tiny particles, then, obviously, such particles move along straight lines. Well, but Earth attracts all material objects – we know that the path of any projectile is curved downward, so the paths of "light particles" we observe also should deviate from straight lines! But Newton did find an answer to that argument. Namely, it is known that the curvature of the projectile's path depend on its speed – the faster the projectile flies, the less its path curves downwards (every sniper or artillery man will testify that this is true!). So, Newton reasoned, if the speed of "light particles" is very high, than one will not be able to observe any deviation from a straigh-line motion. Newton had no clear idea of the value of the speed of light – however, at the times he lived it was already known that it was a very high speed.

As argued by Newton, the law of reflection – namely, that the angle of reflection is equal to the angle of incidence – can also be explained by the laws of mechanics. Consider a rubber ball dropped vertically on the floor: it "bounces back", but it loses some of its speed. Balls made of special plastic lose less speed, and a steel ball bouncing back from a wall of solid steel may lose as little as 1% of its speed. If the ball loses no speed at all, we call it a "perfectly elastic collision". So, according to Newton's explanation, the reflection of "light particles" from a mirror – if the light hits the mirror at a right angle – is a "perfectly elastic" process, the "light particles" are "backreflected" with the speed exactly equal to their speed before the collision. OK, but what if the path of "light particles" is not along the "normal" (*normal* in geometry is a line perpendicular to a plane), but makes an angle θ with the "normal"? Well, then, the Newton's reasoning is the following. Suppose that the velocity of incoming "light particles" is V. Velocity in physics is a vector – we indicate that by adding a small arrow symbol at the top of V, so it becomes $\vec{V}$. Now, physics allows us to "decompose" any vector into two perpendicular components. Let's make one of those components parallel to the "normal": we call it V_x, and its value is $V_x = V\cos\theta$. The other component, call it V_y, is perpendicular to the normal, and its value is $V_y = V\sin\theta$.

If a vector of length V is decomposed into two perpendicular components, V_x and V_y, then, according to the Pythagorean Theorem, it must be:

$$V_x^2 + V_y^2 = V^2$$

Is it so with the two components we have chosen? Let's check:

$$V_x^2 + V_y^2 = V^2\cos^2\theta + V^2\sin^2\theta = V^2(\cos^2\theta + \sin^2\theta) = V^2,$$

because, according to the well-known trigonometric identity:

$$\cos^2\theta + \sin^2\theta = 1,$$

for any θ angle. So, the way we have decomposed the $\vec{V}$ vector is correct.

Now, what happens if the "light particle" collides with the mirror? Well, the normal component of the velocity V_x will behave as in the case of a perpendicular collision, i.e., the light particle will be "bounced back" with the same speed, so that V_x changes to $-V_x$.

And what with V_y? Would it change? Let's think: what is needed to change the velocity of an object? The answer is given by the Newton's *Second Law of Dynamics*:

$$F = m \cdot a, \quad \text{or} \quad a = \frac{F}{m},$$

where F is the force acting on an object, m is the object's mass, and a is the acceleration resulting from the application of force. And acceleration, if we use a simpler language, is a *change of velocity over time*. In conclusion: if there is no force, there is no change in the velocity. And there is no force of any kind that act **parallel** to the mirror surface! So, the parallel velocity component remains unchanged in the process of reflection. In other words, before the reflection the velocity components are V_x and V_y, and after the reflection they are $V_x' = -V_x$, and $V_y' = V_y$, as shown in Fig.1.1.

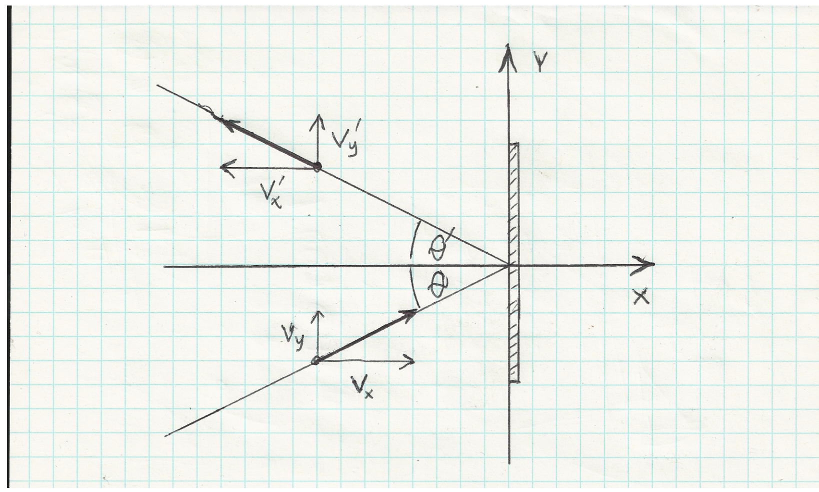

Figure 1.1: The reflection of light from a mirror, as explained in the Newton's theory of "light particles". The normal velocity component V_x changes its sign as the effect of a "perfectly elastic collision" with the mirror surface, and the parallel component V_y remains unchanged.

From the figure it can be readily seen that the tangent of the incident angle θ is $\tan\theta = V_y/V_x$, and the tangent of the "angle of reflection" θ' is:

$$\tan\theta' = \frac{V_y'}{V_x'} = \frac{V_y}{-V_x} = -\tan\theta$$

From the general properties of the tangent function, we know that for any angle φ:

$$-\tan\varphi = \tan(-\varphi)$$

So, as follows from the above, the angles θ and θ' are of equal magnitude – and the $\theta' = \theta$ result we get from the geometric considerations simply means that the the paths along which a "light particle" moves, respectively, before and after the act of reflection, are *symmetric with respect to the normal to the reflecting plane* (i.e., the x axis) – in perfect agreement with the behavior of "real light" in a "real process of reflection" from a mirror plane.

Refraction – let's recall how it works in most simple cases, e.g., when light is incident on the surface of water. Some light is always reflected from the surface, but we have already discussed reflection – now we are interested in the light that passes trough the surface and propagates through the body of water – as shown in Fig. 1.2.

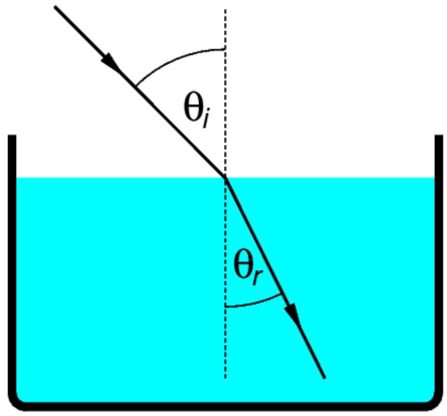

Figure 1.2: Refraction: the path of a beam of light, after entering water, is “bent downwards”. The angle between the normal to the water surface and the incident beam direction is θ_i, and the angle between the normal and the refracted beam direction is θ_r.

The explanation of refraction in the Newton’s theory is also based on relatively simple velocity considerations. The underlying assumption is that “light particles” have energy that consists of two components: the kinetic energy K, and potential energy U. The total particle energy is conserved, meaning that $K + U = \text{const}$. However, the potential energy of the light particle depends on the medium the light passes through: the energy U is different in air, different in water, different in glass, etc.. Therefore, when light passes from one medium to another, the kinetic energy K also changes in order to maintain the total energy constant. And since the kinetic energy of the “light particle” depends on its velocity V as $K = mV^2/2$, the velocity also changes with each such passage.

Newton did not explain what’s the origin of the potential energy of “light particles” – his assumption was what we call a *phenomenological one*: you make the assumption, not necessarily specifying why do you think that “it is so”. But if your assumption leads to a result consistent with the observed phenomenon, you may claim: *I don’t know exactly why things are such as I have assumed – but since the assumption leads to a result consistent with the physical reality, then they must be correct!*

Well, sometimes they indeed are – but not always. So, was the Newton’s assumption correct? Well, in order to answer the question, we have to examine the theory in closer detail.

In fact, a good way of doing that is to use a “mechanical model” – we often do that, because mechanical models are usually "intuitively clear". So, in the present case, consider a "deck" of the type that is popular in Corvallis. Let the deck consist of two parts, one of them at a slightly higher level (e.g., with $/DeltaH$ of six inches, or so) than the other. But let the "inteface" between the two parts has not the form of a sharp step, but of a "step with rounded edges", as depicted in Fig. 1.3. Why rounded edges? Well, because we will be rolling balls on the "double deck", and over the step – and we want the balls never to lose contact with the deck's surface. In addition, let's assume, that friction effects in the ball's rolling motion are negligibly small.

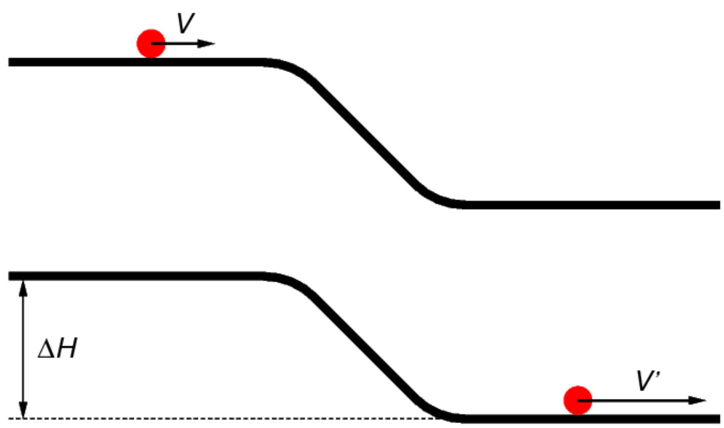

Figure 1.3: The profile of a "double deck" with a "rounded step" betwen the two flat parts. Upper plot: a ball rolls on the upper deck with a speed of V. Lower plot: after rolling down the rounded step, the ball's potential energy decreases by $\Delta U = mg\Delta H$. But the total energy is conserved, so the balls kinetic energy K increases by the same amount – and, consequently, the ball's speed increases.

In Fig. 1.4 there is a view from above at a ball rolling on the deck. But now the ball's motion direction is not perpendicular to the step, but makes an angle θ_i with the perpendicular line. So, the ball now has two velocity components: V_x pointing towards the step, and V_y parallel to the step. Now, when the ball rolls down the step, only V_x increases and at the lower deck it takes the value of $V_x' > V_x$ – while V_y remains all the time the same, because there is no force pushing it in a direction *parallel* to the step. Consequently, the angle θ_r 0n the lower deck is smaller than θ_i on the upper deck – in analogy to the situation shown in Fig. 1.2, when a light beam is refracted by water.

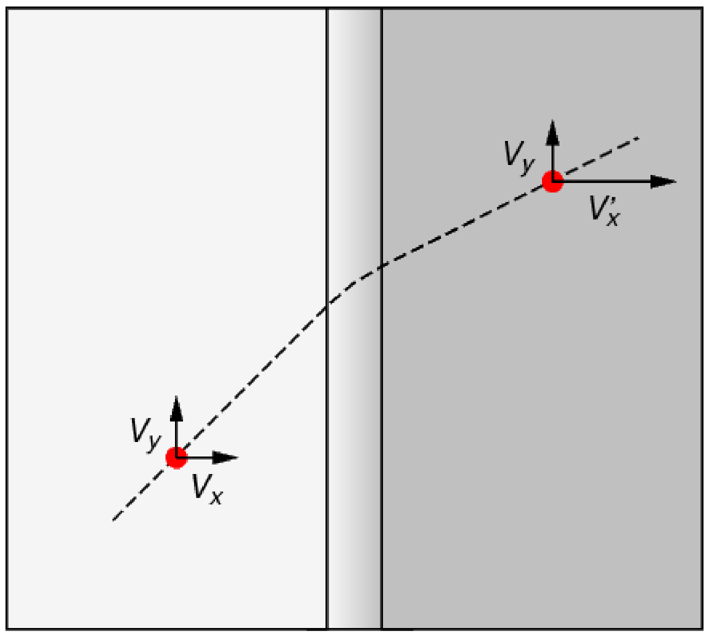

Figure 1.4: A ball rolling over the "double deck", as seen from above. The lighter gray part is the higher one, the darker gray is the lower part, and the "rounded step" is where the gray gradually changes from lighter to darker. When the ball starts rolling down, only the V_x component increases. There is no force acting parallel to the "step", so that the V_y velocity component remains unchanged. It causes that at at the lower deck's part the ball rolls along a line that makes an angle with the direction perpendicular to the "step"(i.e., the x direction) which is smaller the angle between the rolling direction and the x axis at the upper deck.

So, the mechanical model seems to work pretty well. One thing may not look O.K. at the first glance – the "step" in the model cannot be too narrow, and in the case of light entering water, the "step" width, i.e., the width of the interface region between air and water is probably comparable to the size of a single water molecule (about 1 nanometer). Well, but if we assume that the size of"light particles" is even smaller than that of a single molecule, the problem disappears.

As you can see, I (Dr. Tom) am trying to defend Newton's theory of light!

But I still have not used all "defensive weapon" that the theory offers. Namely, the agreement of the Newton's theory with refraction of real light in water (or another transparent medium) is not only qualitative – the mechanical model also obeys an important law of optics, known as the *Snell's Law* (in Europe, it's often called the *Snellius Law* – "Snellius" is the Latin version of the same name). The Snall's Law states that when light passes from one medium to another, the ratio of the sine of the angle of incidence (θ_i – see Fig. 1.2) to the sine of the angle of refraction is always the same – no matter of what the angle of incidence is:

$$\frac{\sin\theta_i}{\sin\theta_r} = \text{constant} = n \tag{1.1}$$

Here, n is a coefficient called *the refractive index*, a real V.I.P. in optics (V.I.P. = very important parameter). Its value depends on the two media involved (e.g. for light impinging water from air, $n = 1.33$; for light entering a diamond, $n = 2.417$, and for light entering diamond from water $n = 1.812$. But right now we are less interested with the values n may take, and more in the fact that n remains constant, no matter what the value of θ_i is.

You may believe me that the Newton's model does obey the Snell's Law – but I don't like to be one of those persons who repeatedly exclaim: *Believe me! Believe me! Bleieve me that it is so!*. Therefore, below I do present a proof – but you don't need to read it line by line, it's enough if... you believe me that it is so :o)). It's not a material that you should know to pass an exam in the Ph332 Course!

***** The proof – you may skip it! *****

For our consideration, we will use Fig. 1.5, which is a slightly modified Fig. 1.4: several symbols have been added, and the image of the "step" is reduced to a single line. The total velocity vector of the ball on the "upper deck" is:

$$V_{\text{tot}} = \sqrt{V_x^2 + V_y^2} \tag{1.2}$$

If m is the mass of the "light particle", then its kinetic energy K is:

$$K = \frac{mV_{\text{tot}}^2}{2} = \frac{m(V_x^2 + V_y^2)}{2} \tag{1.3}$$

The kinetic energy K' of the ball after it runs down the "step" is:

$$K' = \frac{mV'^{\,2}_{\text{tot}}}{2} = \frac{m(V'^{\,2}_x + V_y^2)}{2} \tag{1.4}$$

Note that V_y at the "lower deck" is the same as that on the "upper deck"– as we have discussed before.

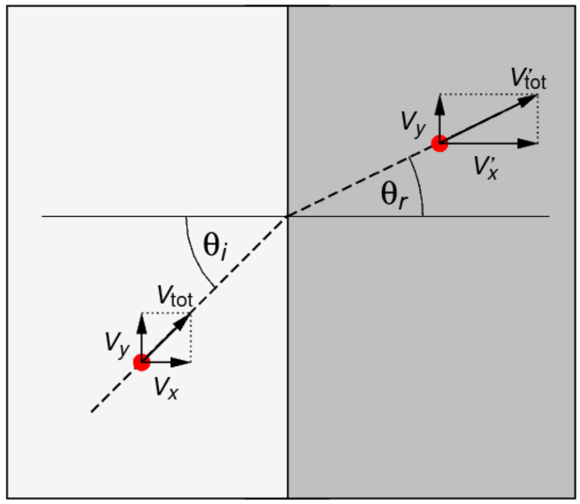

Figure 1.5:

The potential energy of the ball on the upper deck, relative to the lower deck, is

$$\Delta U = mg\Delta H \tag{1.5}$$

(see Fig. 1.3). This energy is converted to extra kinetic energy on the lower deck. Therefore, the kinetic energy K' of the ball on the lower deck is:

$$K' = K + \Delta U \tag{1.6}$$

Now, if we insert into the Equation 1.6 the expressions from the three preceding Equations, we obtain:

$$\frac{m(V_x'^{\,2} + V_y^2)}{2} = \frac{m(V_x^2 + V_y^2)}{2} + mg\Delta H \tag{1.7}$$

Which simplifies to:

$$V_x'^{\,2} + V_y^2 = V_x^2 + V_y^2 + 2g\Delta H, \tag{1.8}$$

and even more:

$$V_x'^{\,2} = V_x^2 + 2g\Delta H. \tag{1.9}$$

We can use this result to obtain an expression for the total speed of the ball on the lower deck, with V'_x eliminated:

$$V'_{\text{tot}} = \sqrt{V_x'^{\,2} + V_y} = \sqrt{V_x^2 + 2g\Delta H + V_y^2} \tag{1.10}$$

And the expression for $V_{\text{tot}} = \sqrt{V_x^2 + V_y^2}$ we got in the Eq. 1.2.

Nowe, we have enough formulae to readily obtain $\sin(\theta_i)$ and $\sin(\theta_r$. From Fig. 1.5 we find that:

$$\sin\theta_i = \frac{V_y}{V_{\text{tot}}} = \frac{V_y}{\sqrt{V_x^2 + V_y^2}}, \tag{1.11}$$

and for $\sin\theta_r$ we get:

$$\sin\theta_r = \frac{V_y}{V'_{\text{tot}}} = \frac{V_y}{\sqrt{V_x^2 + V_y^2 + 2g\Delta H}} \tag{1.12}$$

Now, we can calculate the ratio of the sinuses we wanted to get:

$$\frac{\sin\theta_i}{\sin\theta_r} = \frac{\sqrt{V_x^2 + V_y^2 + 2g\Delta H}}{\sqrt{V_x^2 + V_y^2}} = \sqrt{1 + \frac{2g\Delta H}{V_{\text{tot}}}} \tag{1.13}$$

It can be rewritten in a more compact form:

$$\frac{\sin\theta_i}{\sin\theta_r} = \sqrt{1 + \frac{gm\Delta H}{mV_{\text{tot}}^2/2}} = \sqrt{1 + \frac{\Delta U}{K}} \tag{1.14}$$

The rightmost term in the above equation shows that for a given ΔU, for all "light particles" of the same kinetic energy K the $\sin\theta_i \sin\theta_r$ ratio has the same constant value, not depending on the angle of incidence θ_i – i.e., the Newton's theory of refraction **is** consistent with the Snell Law, **Q.E.D.**

***** End of the "extra" material *****

Newton also worked out an explanation for light diffraction effects – well, from today's perspective we can say it was pretty awkward, and even then interested scientist did not consider it very elegant. However, the successful quantitative explanation of light refraction, together with the great splendor Newton has earned by his discoveries of Laws of Dynamics, and the law of universal gravitation, were enough – Newton's theory was widely accepted, whereas the Huyghens theory of light waves was soon almost forgotten. It did not help that Huyghens theory of wave-like light nature **also was fully consistent with the Snell Law (we will discuss in detail somewhat later**.

Anyway, the "mortal blow" for the theory of "light particles" was delivered a century (97 years, to be exact) after the publication in 1704 of Newton's great treatise on optics (titled Opticks: or, A Treatise of the Reflexions, Refractions, Inflexions and Colours of Light – you may read more about that famous Newton's book from a Wikipedia article, if you click on the blue title provided). The "executioner" was a Britsh 0 scientist Thomas Young, who carried out an experiment, today widely accepted as one of the most important experiments in physics of all times. Namely, he observed that if sunlight passes through a a system of two narrow parallel slits, it forms a pattern of bright and dark "fringes" on a screen placed in a darkened room behind the slits (we will discuss the Young's experiment in greater detail somewhat later, but if you wish to learn more about it right now, please click on this link to an Encyclopedia Britannica article, an watch the movie enclosed – or on this link to a good YouTube short movie).

Such an effect could only be explained on the grounds of the Huygens wave theory. Huygens was vindicated, and the Newtons theory of "light particle" was "pronounced dead".

The period of nearly 90 years which followed, was a series of great triumphs of the wave theory of light. On its ground Fresnel and Fraunhofer worked out elegant theoretical approaches to light diffraction (see this Hyperphysics entry for details) – calculations based on their theories were found to be in excellent agreement with experimental observations. Another great achievement was the explanation of the phenomenon of light polarization. In 1849 Armand-Hippolyte-Louis Fizeau, a highly talented French experimenter, performed the first "on-Earth" measurements of speed of light c[1].

[1] Earlier determination of c had been done by astronomical methods – one of them was based on the Ole Roemer's idea of determining c by observing anomalies in the eclipse time of Jupiter's moon Io (a common textbook error is that Roemer had determined c to be 131 000 miles/second – Roemer could not present any number, because he **did not know the Earth's orbit diameter** – it was determnied with nearly-perfect accuracy only from observations of Venus transits across the Sun's face in 1761 and 1769). In 1727 an English astronomer, James Bradley, discovered the phenomenon of stellar aberation, i.e., an apparent shift of a star position due to the orbital motion of Eart (see the second picture in the linked page). The c value determined by Bradley is given in many texbooks as 301,000 km/s – again, it's probably a mistake, Bradley did not know the orbital speed

The "on-Earth" result was important, because the the earlier determinations, obtained by astronomical methods, and not by using a "hands-on" apparatus, were not universally trusted. Fizeau's results obtained by the tooth-wheel method, and shortly afterwards a more accurate result obtained by a rotating-mirror method developed by Leon Foucault, were fully trusted. What more, Fizeau and Foucault were able to directly determine the speed of light in water, showing that it was slower than in vacuum – in agreement with the Huyghens' wave theory, but in striking disagreement with Newton's "corpuscular" theory, delivering one more mortal blow to the latter.

So, after all such developments nobody had doubts anymore that light has a wave-like nature. Still, it was not clear what was undulating in the light wave? In all other known form of waves, there always "was somtething" to vibrate: in sound waves, it was the air density; in waves on water, it was water oscillating up and down; and so on. But in light? It was easy to check that light was passing without any problem through vacuum. Vacuum is an empty space, what can vibrate in it? It seemed to be something magic...

The answer to that riddle was found by a Scottish scientist James Clerk Maxwell. His great achievement was to unify in 1861-62 the theories of electric field and magnetic field. Written in "mathematical language", the theory has the form of four differential equations (soon scientists started calling them the *Maxwell Equations*). Differential equations differ from algebraic equations in that their solutions are not numbers, but functions – and often sets of differential equations have more than one possible solutions. When examining his equations, Maxwell discovered that one of the possible solution has the form of a function describing a propagating wave, the so called "wave function". He called it an "electromagnetic (EM) wave". And from the wave equation one can figure out the speed of the wave's propagation – the speed Maxwell obtained from his calculations was very close to the measured speed of light c.

So, it seemed that the mystery of the light's nature was finally solved! Only one thing was missing, the Maxwell EM waves still existed only on paper. A proof was necessary that the EM waves have a real existence. The only way of making light in the middle of the 19th century was from bodies heated up to high temperatures (e.g., heated gases such as candle or torch flames, molten iron, etc.). But was the glow from a hot body really an EM wave? It was completely unclear how Maxwell Equations could be

of Earth, so he could not obtain a result that accurate.

used for describing the phenomena occurring in hot bodies (in fact, physical phenomena responsible for the emission of light by heated bodies became clearly understood only in the 20th Century).

In Maxwells days, a sufficient proof for the existence of EM waves would be if someone could create an such waves using an apparatus, the action of which would be consistent with the conditions necessary for generating EM waves, predicted by the Maxwell Equations. And such an apparatus was successfully build by in 1886 by a German accomplished experimental physicist, Heinrich Hertz. Guided by the Maxwell's theory, he build devices that from today's perspective can be called a "simple radio transmitter" and a "very simple radio receiver". And he was able to send signals over a distance from the transmitter to the receiver! He also performed thorough studies of what carried the signal from the transmitter to the receiver, and showed that it had all properties of an EM wave predicted by the Maxwell's theory (Regretfully, Maxwell did not live to see that moment of the great triumph of his "brainchild" – he died in 1879). Hertz waves were not light – their wavelengths were meters, million times longer than the wavelength of visible light. But since all other characteristics of the "Hertz waves" and light appeared to be fully consistent, it was widely accepted that the nature of light had been finally explained.

However, it turns out that Mother Nature has a perverse sense of humor. Namely, Heinrich Hertz, the same person who dispelled last doubt that light is an electromagnetic wave, during his famous experiments fortuitously observed a strange effect caused by light incident on a metal. In short, he discovered that ultraviolet light incident on metal cause "the emission of electric charge" from its surface. Intrigued by Hertz's observations, Alexandr Stoletov, a Russian physicist, conducted first thorough research on that strange effect over the 1887-91 period. His work stimulated undertaking more studies by other scientists – by the end of the 19th century it became clear that light can "knock out" electrons from a metal. The name coined for that mysterious phenomenon was *Photoelectric Effect* (PE).

Such discovery came out as a big embarrassment to scientists. It seemed obvious to everybody that any "wave" has a wide front. It was known how much energy was needed to eject an electron from a metal, and it was clear that the wave must deliver such energy to a small area, of a size comparable to one or a few atoms, in a very short time, because "kicking out electrons from metal" was found to be a very quick process. So, the power of a *continuous* wave capable of ejecting electrons from metal surface would in microseconds cause the metal to melt and evaporate. Yet, the PE was occurring even for a very low light intensity!

The scientist who resolved that "mystery" in 1905 was nobody else than Albert Einstein (in 1921, he was awarded a Nobel Prize for that). One can say that Einstein "resurected" the theory of light particles. But definitely it was not a "return with vengeance" of Newton's light particles. Einstein's "particles", whom he called *light quanta*, were – I'm looking for a good word – objects one can think of as created by "fragmentation" of a wave. Einstein's term "light quanta" did nor survive for very long, somebody started using instead the term *photon*, which was widely accepted and people started using it right away. So, a photon is an individual object that carries a portion of energy – like a particle. But it also retains some typical characteristics of a wave – namely, the wavelength and the frequency. The photon energy (E) and its frequency (the common symbols used are either the Roman f, or the Greek ν) by the famous Einstein's formula:

$$E = hf \quad \text{or} \quad E = h\nu \tag{1.15}$$

where h is the so-called *Planck Constant*, $h = 6.6261 \times 10^{-34}$J·s (note that the frequency unitg is 1/s, so that the energy calculated by the above equation comes out in the right unit, the Joule).

In view of the above, one can hardly say that Einstein "resurrected" Newton's light particles – the latter hardly 'knew" what the wavelength and the frequency was. So, we cannot say that Einstein's theory was a "resurrection" of Newtons theory – a much better term would be *reincarnation*. You know that *reincarnation* is a *new version of something from the past*, but not necessarily of the same kind as the old version. For instance, a sweet person may be a reincarnation of a dove or a lamb, and a person of violent temper may be the reincarnation of a lion or a tiger.

Then, who is wrong and who is right? These who say that light is a wave, or those who say that it has a *corpuscular*, i.e., a particle-like nature? no one is wrong, they are both right. Light simply exhibits a *particle-wave duality*, a dual nature . In a distant analogy to *Dr. Jekyll and Mr. Hyde*[2] from the famous 1886 short story by Robert Louis Stevenson. Dr. Jekyll and Mr. Hyde were the same person, but of completely different personalities – and one could be transformed into the other, or vice versa.

[2]For me, a better title would be *Dr. Jekyll and Mr. O'Hyde*, because in Polish, my native language, the word *Ohyda* means something or someone extremely evil.

In certain physical situations – or, one can say, in certain physical phenomena – light clearly exhibits its wave-like nature: for instance, in diffraction and interference effects, or in such a simple effect as refraction. In photoelectric effect light clearly demonstrates its "corpuscular", i.e., its particle-like nature. Another example of the latter is emission of light by atoms.

We will begin with discussing a number of effects for which a description in terms of the wave theory is appropriate – and only towards the end of the course, we will switch to phenomena in which light clearly demonstrates its particle-like nature. Therefore, Chapter Two will begin with a general description of waves.