Problem № 15: Optical Tunneling

M. Lukas[1], M. Marković[2], D. Pelc[1], M. Popović[1], I. Sudić[1]
[1] *V. gimnazija, Klaićeva 1, 10000 Zagreb, Croatia*
[2] *XV. gimnazija, Jordanovac 8, 10000 Zagreb, Croatia*

The problem

Take two glass prisms separated by a small gap. Investigate under what conditions light incident at angles greater than the critical angle is not totally internally reflected.

Introduction

If light is entering a prism at an angle greater than the critical, the ray will be totally reflected. However, if we place another prism close to the reflection plane of the first prism, some light will make it across the gap between the prisms. This effect, which is in fact the optical analogy of the quantum – mechanical potential barrier tunnelling of particles, is called optical tunnelling. As we shall see, it can be explained both classically with the aid of the Maxwell equations and quantum – mechanically, both explanations giving the same results; but as the classical picture is somewhat closer to us we will use the Maxwell explanation.

In the beginning, before we make any quantitative analyses, we will mention some properties of the tunnelling process observed experimentally; the conclusions we shall draw from the observations will be of great aid in constructing the mathematical model:

- The intensity of the tunnelled light falls off exponentially when the gap is increased linearly
- The passing light can be observed as long as the gap is a few wavelengths wide; if the gap is further increased, the light is too faint to detect, which means that the intensity falls off very rapidly
- The intensity of the tunnelled light depends on the refractive index of the medium between the prisms
- Naturally, the effect cannot be seen if the refractive index of the medium is larger than the refractive index of the prisms because total reflection doesn't occur either

The explanation of all these effects is in fact very simple: when the wave is reflected off the prism surface, the electromagnetic field can't be discontinuous at the boundary between the prism and the medium beyond, it has to extend a little further into the medium, decaying rapidly. That field can indeed be detected, and in the classical picture it presents the smooth transition between the field in the prism and the ″no – field″ in the medium beyond the reflection plane. The most understandable explanation of the occurrence of this transition uses Huygens' principle; it is known that a light wave in a crystal (here the prism) or medium is the result of the interference of all waves scattered on the atoms of the medium. During total reflection there is only one principal interference maximum, the reflected ray. However, in the vicinity of the reflective surface ″tails″ of the light that didn't manage to interfere completely are formed, decaying fast. These tails are referred to as the evanescent wave. If the second prism is placed in that region a new wave will be formed in that prism because of the ″tails″ shaking its atoms generating secondary emissions. Only in that case does energy leave the first prism; the evanescent wave itself carries no energy whatsoever because the electric and magnetic fields are in counterphase. The more beautiful quantum – mechanical explanation uses the Indeterminacy Principle: a photon cannot be localized with 100% accuracy, so there is always a finite probability that some of them are beyond the reflection surface, in the ″forbidden″ region. The classical, as well as the quantum theory, gives very reliable and simple mathematical results, although the interpretations are somewhat different. As we already mentioned, we will follow the classical theory, using the boundary conditions for the light waves on the reflection surface[1,2]. In the first part of the article we will present a short quantitative description of this theory and the determination of the dependence of the tunnelled light intensity on the main parameter – the gap width – in order to proceed to the experimental results and their evaluation. The quantum – mechanical theory won't be examined in detail because it gives the same numerical results as the classical[3,4].

Theory

To gain an exact relation between the tunnelled light intensity, the gap width and the refractive index of the medium between the prisms, we have to solve the Maxwell equations implementing the boundary conditions for reflection. As the prism material and the medium between them are dielectrics, we will work with Maxwell equations for a dielectric medium:

$$\nabla \mathbf{E} = -\frac{1}{\varepsilon_0}\nabla \mathbf{P}$$

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}$$

$$\nabla \mathbf{B} = 0$$

$$c^2 \nabla \times \mathbf{B} = \frac{1}{\varepsilon_0}\frac{\partial \mathbf{P}}{\partial t} + \frac{\partial \mathbf{E}}{\partial t}$$

where **E** is the electric field, **P** the polarization, **B** the magnetic field induction, ε_0 the permeability of vacuum and c the speed of light *in vacuum*. We know that plane sinusoidal waves of the electric and magnetic fields are particular solutions of these equations; the electric field is

$$\mathbf{E} = \mathbf{E}_0 e^{i(\omega t - \mathbf{kr})}$$

with $\mathbf{E}_0$ the amplitude, ω the circular frequency, t time, $\mathbf{k}$ the wavenumber vector (which points in the direction of wave propagation and has a length equal to $\frac{\omega n}{c}$, with n being the refractive index of the medium through which the wave is propagated) and $\mathbf{r}$ a radius vector. For the magnetic field one readily obtains

$$\mathbf{B} = \frac{1}{\omega}\mathbf{k} \times \mathbf{E}$$

The magnetic field is thus perpendicular to the electric field, a well – known result. What concerns us now are the amplitudes of the fields in the three regions: the first prism, the gap, the second prism. To obtain them we must find the boundary conditions at the region boundaries. First we must define the geometry and polarization of the incident light, because the relation between the direction of the electric field vector and the plane of propagation is quite important. The coordinate system will be such that light propagates in the xy – plane, and as a special case we will take the light to be linearly polarized in the z – direction. The reflecting plane of the prism lies in the yz – plane, and due to the ray being practically infinitely thin we assume the face to be of infinite size. The face of the second prism is parallel to the first, the distance between them being d. The wave vector of the incident light we have denoted by $\mathbf{k}_1$, the vector of the reflected light is $\mathbf{k}_r$ and the vector of the transmitted light $\mathbf{k}_t$. However, as we shall see, this vector is complex in the space between the prisms, becoming real in the second prism; that real vector is denoted by $\mathbf{k}_t'$. The notation is analogous for the electric and magnetic fields. The angle of incidence or reflection (it is clear that those two angles are equal) is φ (Fig. 1). Now we can write the waves in the first prism in the exponential form;

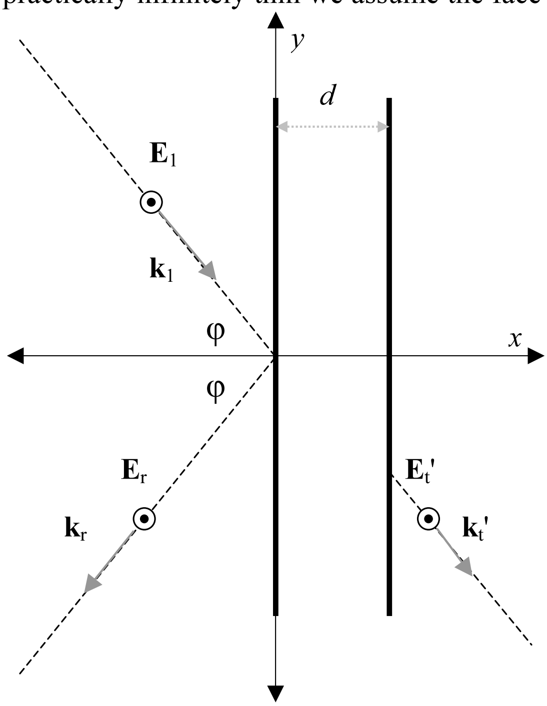

Fig. 1. Geometry of the problem

$$E_1 = E_{10} e^{i(\omega t - k_{1x}x - k_{1y}y)}$$

is the incident electric field, and

$$E_r = E_{r0} e^{i(\omega t - k_{rx}x - k_{ry}y)}$$

the reflected field. In the space between the prisms the wave will be

$$E_t = E_{t0} e^{i(\omega t - k_{tx}x - k_{ty}y)}$$

Now we want to find the relations between the wave vector components of the incident, reflected and transmitted waves. Due to energy conservation we can readily conclude that

$$E_1 + E_r = E_t$$

for all times and coordinates; that induces:

- The frequencies of the incident, transmitted and reflected waves are the same
- For the amplitudes we get

  $$E_{10} + E_{r0} = E_{t0}$$

  and

  $$k_{1x}E_{10} + k_{rx}E_{r0} = k_{tx}E_{t0}$$

  because the derivatives have to be continuous at the boundary (or the wave equation would break down at the edges which is physically impossible). Knowing that $k_{1x} = -k_{rx}$ (the incident angle is equal to the reflected) we obtain

  $$E_{r0} = \frac{k_{1x} - k_{tx}}{k_{1x} + k_{tx}} E_{10}$$

  $$E_{t0} = \frac{2k_{1x}}{k_{1x} + k_{tx}} E_{10}$$

  Of course, if the wave vector of the transmitted wave is complex, the amplitude is equal to the modulus of the complex amplitude obtained with the above relation.
- Due to light speed constancy, one obtains for the wavenumbers of the incident and transmitted waves

  $$k_{ty} = k_{1y}$$

  $$k_{tx} = \left(\frac{n_1}{n_0}\right)^2 k_1^2 - k_{1y}^2$$

  where n_0 is the refractive index of the prisms and n_1 the refractive index of the medium between them.

Now we are in possession of all ingredients necessary for finding the equation of the wave in the gap. For total reflection, the angle φ must be greater than the critical angle defined by

$$\frac{n_0}{n_1}\sin\varphi_c = 1$$

where φ_c is the critical angle. That relation is of course just a special case of Snell's law. As $k_{1y} = k_1 \sin\varphi$, the x – component of the wave vector of the transmitted wave becomes

$$k_{tx}^2 = k_1^2\left[\left(\frac{n_1}{n_0}\right)^2 - \sin^2\varphi\right]$$

If the angle of incidence is larger than the critical angle, $\sin^2\varphi > \left(\frac{n_1}{n_0}\right)^2$ and the x – component of the wave vector is a pure imaginary; inserting it into the wave equation one gets the predicted exponential drop of amplitude (the positive solution, corresponding to exponential growth, makes no physical sense):

$$E_t = E_{t0}e^{-\frac{n_1}{n_0}k_1 x\sqrt{\left(\frac{n_0}{n_1}\right)^2\sin^2\varphi - 1}}e^{i(\omega t - k_{1y}y)}$$

If there is only a medium of refraction index n_1 the field in this medium will decay very fast and the reflected wave will show no energy losses; energy only oscillates to and fro at the boundary but doesn't leak into the medium. That can be easily shown considering the Poynting vector; the cross – product (and accordingly the vector itself, which represents the energy carried by the wave) of the electric and magnetic fields is zero duo to their counterphase oscillations. However, if we put a second prism near to the first, the evanescent wave will shake the electrons in the second prism and cause another emission, with the fields in phase this time, and carrying energy. The energy is taken from the reflected wave, causing it to faint as the prisms are drawn closer, vanishing when they touch. Following these arguments (or referring to the expressions for the amplitudes found above) we arrive at the formula for the wave intensity in the second prism:

$$E_t' = E_{10}e^{-2\pi n_1\frac{d}{\lambda}\sqrt{\left(\frac{n_0}{n_1}\right)^2\sin^2\varphi - 1}}$$

with λ the wavelength *in vacuum.* Due to simplicity we introduce a parameter Θ:

$$\Theta = 2\pi n_1\sqrt{\left(\frac{n_0}{n_1}\right)^2\sin^2\varphi - 1}$$

which will be measured and compared to theoretical values. To sum it all up, the presented Maxwell theory arrived at an easily investigable relation between the intensity of the tunnelled wave and the relevant parameters (gap width, refraction indices) considering only boundary conditions at the interfaces. As we shall see, the obtained expression also quite well agrees with the experimental data we produced.

Experiment

In order to obtain a larger precision and reliability, the measurements were conducted in two different frequency ranges: microwaves and visible light. With the microwaves the investigation of the phenomenon posed no problem because the gap can be up to a few centimetres wide; however, the procedure is more complicated in the optical range due to the very small gaps necessary for the effect to be measurable. The prisms must be very smooth and clean, and the measurement of the gap size is rather difficult[5]. That's why we conducted the measurements in time – the intensities of the tunnelled light were logged at fixed time intervals while the prisms were uniformly moving towards each other pushed by a slow motor. That method gives quite good results but the fast logging necessary for such small distances (a few microns!) can be a problem. But using a very slow motor (1/2 rpm) and large sampling we succeeded in measuring the effect.

The setups for the optical and the microwaves experiments look rather similar (Fig. 2.). The prisms for the microwaves were made of paraffin, with dimensions 150x150x100 mm. The gap between the prisms was varied with a simple wooden translator moved by a long screw. The source of the necessary radiation was a

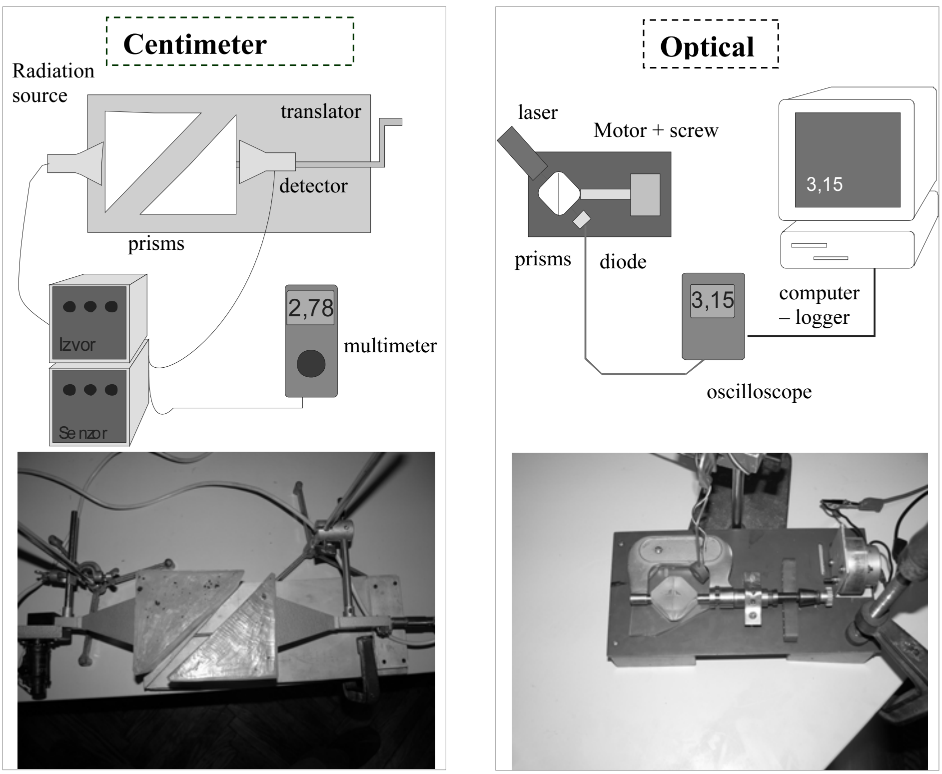

Fig. 2. Experimental setups. The setup for microwaves is on the left and the optical setup on the right

magnetron pentode with an amplifying horn which was placed close to the prism during measurement. The detector for the radiation which was used in tunnelled wave measurement consisted of an amplifying horn with a resonant space containing a 100mH coil which was placed at a node of the wave in the tube. Thus the voltage read from the coil is directly proportional to the field, not the intensity (like when working with photodiodes). The wavelength of the waves was 3.0 cm, making little scratches and defects on the prisms unnoticeable. The optical setup was quite similar, only that the prisms were made of glass, mounted in the screw driven by the motor. The velocity of the prisms was 0.025 cm/min (0.6 seconds for a micron), and the intensity was logged every fiftieth of a second. The data had later to be smoothed due to imperfections (and dust) on the prisms and noise, resulting in a 10 measurements per second sampling. A 780 nm laser diode with polarizer was used as the light source. The transmitted light was detected by a photodiode which gave signal proportional to the intensity (the *square* of the electric field) of the light.

The results of the measurements are in relatively good agreement with the theoretically predicted curves, the agreement being better for the microwaves due to much greater precision and a larger number of points (Fig. 3.). The slight deviations from the theoretical curve at low intensities are probably caused by the unfocusedness of the source or voltage variations on it; anyway, the deviations were completely random throughout the measurements so they are probably some indeterminate noise. In the optical range (Fig. 4.) the disturbances were very large, making high – sampling measurements impossible, and the discrepancies are larger due to smaller precision and defects (for example dust particles, the size of about a micron, can influence the results badly).

A complete comparison between experiment and theory wasn't possible because we didn't know the relation between the measured detector voltage and the real field, but as the functions look the same it was enough to compare the decaying factors in the exponents. In the optical case we have to perform a little transformation due to the time – measurements; the gap width, d, becomes $d_0 - vt$, with d_0 being some starting value

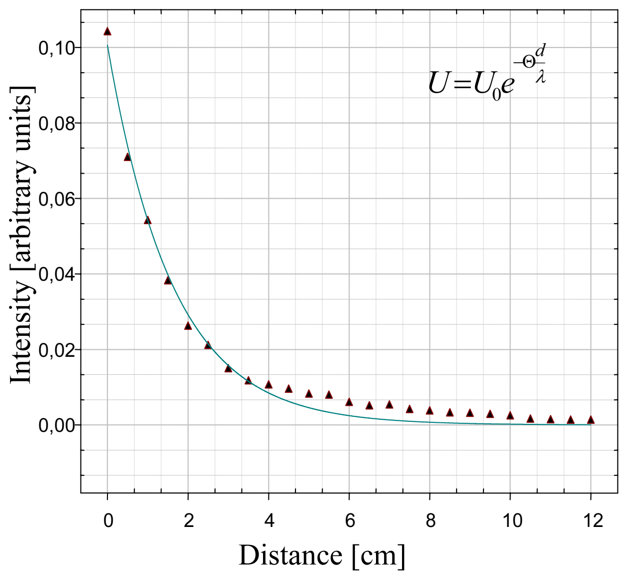

Fig. 3. The experimental curve and fit for the microwave experiment

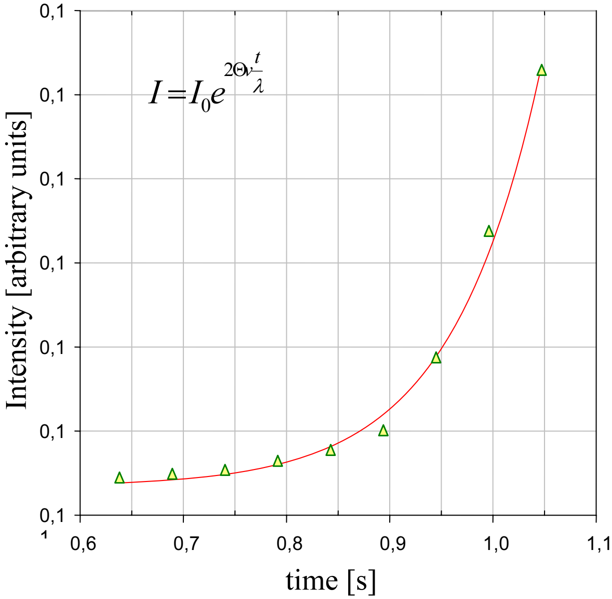

Fig. 4. The experimental curve and fit for the optical experiment in time mode.

(not necessary for the comparison), v the prism velocity and t time, the measured variable, and of course we had to account for the fact that the measured output voltage of the photodiode was propotional to the square of the field. The refractive index of the glass prisms was 1.48 (measured) and the paraffin had 1.50 for our wavelength. The refractive index of the air between the prisms was taken to be 1.00 and the angle of incidence was 45˚ . That leads to the following results for the decay factor Θ (all the data have been given again for sum – up):

Prism refractive index, n_0	1.48
Wavelength	780 nm
Angle of incidence	45˚
Theoretical Θ	1.9
Experimental Θ	1.1±0.1

Prism refractive index, n_0	1.50
Wavelength	3.0 cm
Angle of incidence	45˚
Theoretical Θ	2.2
Experimental Θ	2.5±0.1

We see that the agreement is rather good, especially in the centimetre range; the measurement was, as mentioned, rather more precise there; the agreement to a factor of 1.7 in the optical range is quite satisfactory considered all errors.

Conclusion

To sum things up, we can conclude that we have approached the problem of optical tunnelling and solved it to a certain depth, appropriate to our resources. We have obtained the fundamental relations between the intensity of the tunnelled light and the gap width using elementary classical theory and checked the theoretical results in experiment, with two different wavelengths in two different ranges of the spectrum. The microwave measurements enabled us to perform high – precision tunnelling measurements and obtain valuable quantitative data thanks to the macroscopical size of the gap, while the optical experiment served a more demonstrational purpose due to the difficulty of obtaining reliable measurements, though we tried to do that too. The parameters entering our intensity formula are the gap width as the most obvious and the prism and medium between prisms refractive indices. In our investigations we mainly focused oh the gap, somewhat neglecting the refraction indices. However, we are of the opinion that the gap is indeed much more important and maybe more fundamental to the effect itself than the refraction indices – their variation in fact only induces a change of the critical incidence angle which is not as important. On the other hand, measurements with varying refractive indices could provide an even more firm confirmation of the theory. Also, our formulas were obtained using classical theory; a more complete investigation of optical tunnelling might include a more detailed treatment of the quantum, photonic picture, linking it to the classical formalism. We have chosen the classical theory in our work due to its simplicity, in spite of the beauty of the quantum model, not having the space or time to make more detailed theoretical investigations.

With the suggested adds, for which we didn't have the time or equipment, included, we could say that the problem of optical tunnelling would be rather completely solved. And in the end we may briefly answer the question posed by the problem itself: light incident at angles greater than the critical angle is not totally internally reflected if the second prism is as close to the first as a few wavelengths of the light used.

Acknowledgements

We thank Silvije Vdović, Ticijana Ban and Goran Pichler from the Institute of Physics for help and advice in conducting the optical experiment, the workshop staff of the IF for help with the optical apparatus, Ivica Aviani and Željko Marohnić for discussions, Tihomir Surić for corrections and suggestions in the theory, Hrvoje Mesić for help with the microwave experiment, Želimir Miklić for the magnetron and horns and our mentor Dario Mičić for support and help with everything.

References

[1] D. B. Sivuhin, *Opći kurs fizike*
[2] R. Feynman, *Lectures on Physics*
[3] E. Mertzbacher, *Quantum Mechanics*, J. Wiley & sons, New York, 1961.
[4] G. Joos, I. M. Freeman, *Theoretical Physics*, Dover, New York, 1950.
[5] D. D. Coon, *Counting Photons in the Optical Barrier Penetration Experiment,* Am. J. Phys. **4**, 240, 1965.