8.1. SOLUTION OF BRAZIL

Problem № 15: Optical Tunneling

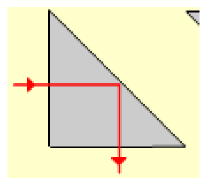
Daniel Nogueira Meirelles de Souza, São Carlos – SP
Escola Educativa – Instituto de Educação e Cultura

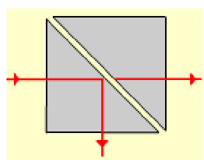
The problem

"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

Total internal reflection is a well known optical phenomenon. It occurs when light propagating in a medium of index of refraction *n* reaches a separation boundary between this medium and one of index of refraction smaller than *n* at an angle of incidence greater than a critical angle θ_c. All light is reflected back into the first medium (of greater index of refraction). Total internal reflection can be well visualized if we have a triangular 90° prism (of, for example, glass) and a light source (for example, a laser pen). If we simply make the light enter the prism and reach the separation boundary with air on its hypotenuse at an angle greater than the critical angle, we will easily see total internal reflection. View top picture.



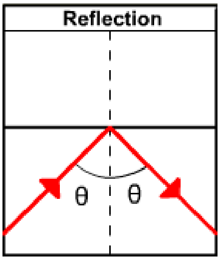


If a second prism or piece of glass is approached to the hypotenuse of the one in which total reflection is taking place, making the two prisms separated by a small gap, an unexpected phenomenon might occur (bottom picture) . A normal total internal reflection would be expected, since air is still surrounding the prism. However, if the gap is sufficiently small, part of the light that would suffer total internal reflection is unexpectedly transmitted into the second prism, leaving a smaller amount of light to suffer reflection:

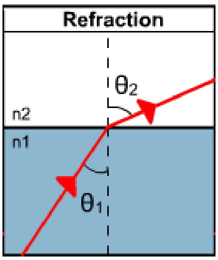
This phenomenon is given the name of Frustrated Total Internal Reflection (which we shall now call FTIR). It was first observed by Isaac Newton, about 300 years ago, and reported in his famous book *Optics*. Newton brought a convex lens close to the region into contact with the reflecting surface of the prism and realized some light started to travel through the lens. Although reported, the phenomenon could not be successfully explained by Newton. In fact, the phenomenon is not predicted by geometric optics. In this problem, we shall investigate FTIR and the conditions in which it occurs.

Theory

Geometric Optics: it is important for the understanding of FTIR knowledge of basic concepts in geometric optics.



Reflection:
Light suffers reflection when it reaches the boundary of a reflecting surface. The angle between the incident ray of light and the direction normal to the reflecting surface is equal to the angle of the reflected ray with this normal. This is the law of reflection.



Refraction:
Is what happens to a wave when it changes its medium of propagation and consequently its propagation speed. In optics, light suffers refraction when it changes its propagation medium into one of different index of refraction. A medium's index of refraction, *n,* is given by:

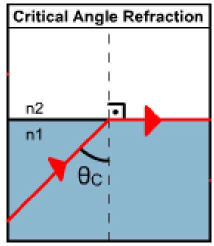
$$n = \frac{c}{v}$$

In which *c* is the speed of light in vacuum (aprox. 3.10^8 m/s) and v is the speed of propagation in the considered medium. The index of refraction is also referred as optical density. When a light ray suffers refraction, its speed and direction changes. The law of Snell – Descartes relates the sine of the incident and refracted angles and the index of refraction of both mediums:

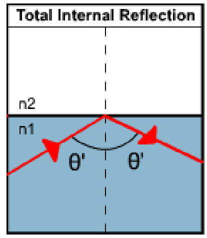
$$\sin \theta_1 . n1 = \sin \theta_2 . n2$$

We can conclude from the equation that if light is traveling initially in a medium of smaller index of refraction and refracts into a medium of greater index, the angle of the refracted ray will be smaller than the angle of the incident ray. If light is traveling initially in a medium of greater index of refraction and refracts into a medium of smaller index, the angle of the refracted ray will be bigger. It is important to add that not all incident light is refracted: part of it can be reflected, returning back into the first medium. If we take the case of the initial medium of propagation be of bigger index, we will find that at a certain angle the refracted ray will be perpendicular to the normal direction and therefore parallel to the surface. The incident angle in which this occurs is the *critical angle*. From Snell – Descartes, we find that the sine of the critical angle θ_C is given by:

$$\sin \theta_C = \frac{n2}{n1}$$

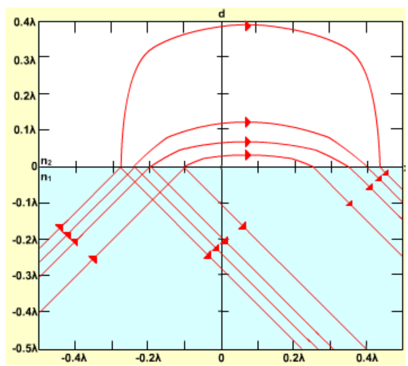


If, in these conditions, the incident angle becomes greater than θ_C, total internal reflection will occur, in which all light is reflected and none is refracted:



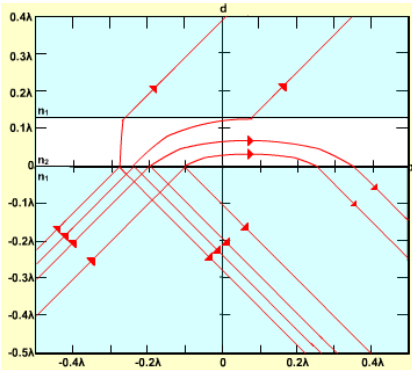
Evanescent wave in Total Internal Reflection:

Frustrated Total Internal Reflection could only be properly explained with 19th century Maxwell's electromagnetic theory. A deeper look into total internal reflection was possible. In total internal reflection, we have the penetration of the electromagnetic wave in the region beyond the totally reflecting interface, into the second medium. This penetrating wave is called the **evanescent wave** (see pic.).



This picture was taken from the third reference. It shows total internal reflection of light incident at 45° on the interface, in which n1 = 1,5 and n2 = 1,0. The flow lines are represented, showing that when light is incident in an angle greater than the critical angle, part of it is reflected back into the first medium and part, surprisingly, actually penetrates into the less optically dense medium, creating an existence of electromagnetic energy in the region beyond the interface, which travels according to the flow lines represented. This is a strange behaviour. The wave that penetrates into the second medium runs along the direction parallel to the interface, and, after a distance of the order of the wavelength λ, returns to the first medium, parallel to the reflected rays. This is what actually happens in total internal reflection. The picture also shows that, increasing the distance d from the interface, we have a smaller concentration of flow lines, which means the amplitude of the evanescent field drops if the distance from the interface is increased, so that, at some distance, the amplitude would be too small to be considered.

If we were to approach a second piece of glass to the first, at a distance in which the amplitude of the evanescent wave is appreciable, we would have that



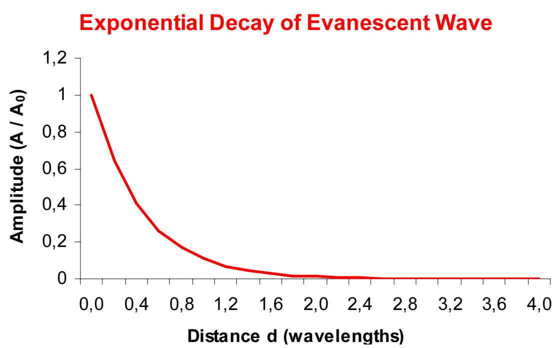
some of the electromagnetic energy would enter the second glass in the form of a light wave. This would result in a smaller amount of light returning to the first medium:

This partially explains frustrated total internal reflection. If the second prism is approached at a distance small enough, it will be able to capture the electromagnetic energy in the evanescent wave. Due to conservation of energy, less light returns to the first prism. If the distance between the prisms is too great, the amplitude of the evanescent wave will be practically zero and no light would be frustrated. The drop of the amplitude of the evanescent wave with the distance from the interface in the direction normal to this interface is exponential:

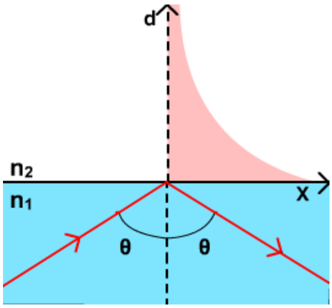
$$A = A_0 e^{-d\alpha}$$

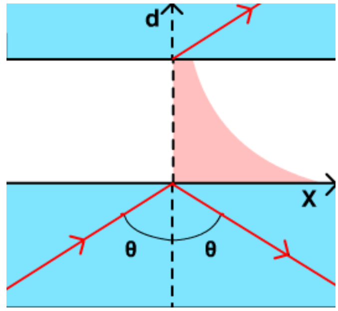
$$\alpha = \frac{2\pi}{\lambda}\sqrt{\left(\frac{n1}{n2}\right)^2 \sin^2\theta_1 - 1}$$

Where A_0 is the amplitude at distance 0, d is the distance in the direction normal to the interface, λ is the wavelength of the electromagnetic wave and θ_1 the incidence angle. The deduction of this equation may be viewed by the curious reader in the **appendix**. From the exponential equation, we can observe that the amplitude of the evanescent wave is considerable at distances smaller to or of the order of λ. The graphic below shows how the amplitude of the evanescent wave (in relation to A_0) varies with distance from the interface (in units of wavelength) when the incidence angle is 45° and n1 and n2 are, respectively, 1,5 and 1:

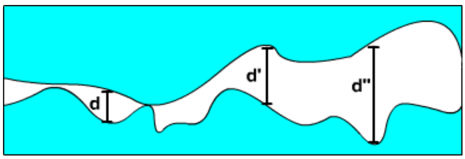


Below we have another representation for the evanescent wave and FTIR, in which the exponential decay is represented by the reddish curve:





So, in order to obtain the phenomenon, we must place our prism at a distance no much greater than 2 wavelengths. The wavelength of visible light varies from approximately 400 to 800 nanometers. We must then place our prisms at a distance of the order of 10^{-7}m from each other. The best way to try to do this is pressing one prism against the other, or at least putting them into maximum contact possible. This is because most prisms are still irregular in microscopic terms.



So we have great differences in distance between the prisms along their surfaces, if compared to the wavelength of visible light. This makes it impossible to obtain FTIR with prisms which don't have a very good surface quality.

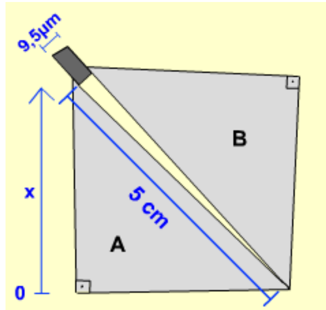
Experiment

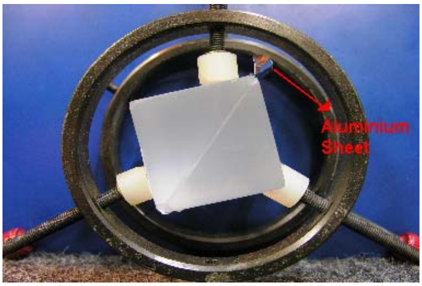
The Objective of the experiment was to verify and measure the dependence of the transmitted light (in frustrated total internal reflection) on the separation air gap d between the prisms.

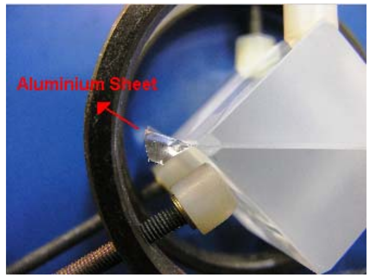
Utilized Materials:

- 2 BK7 45° Prisms (n = 1,515 for 635nm)
- Red laser (λ = 635 nm)
- Aluminium Sheet (9,5 ± 0,5 μm width).
- Photocell
- Voltmeter
- 2 lens
- Micrometer
- 2 Fasteners
- Voltmeter

Experimental Methods:
With the use of the very thin aluminium sheet, the prisms were arranged so that the air gap (or distance) between them changed uniformly. The prisms used had a very good surface regularity:







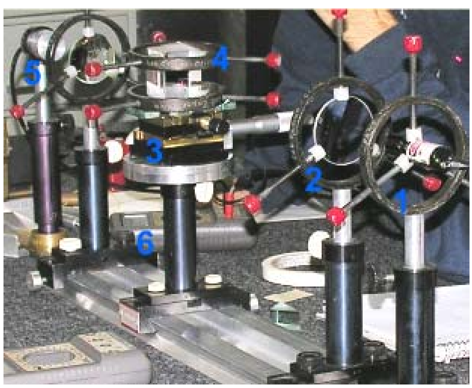
2 Fasteners were used to press the prisms against each other in order to assure they were really in contact at one end. The prisms used had a very high optical quality. Along a distance of 5 cm, the separation gap between the prisms changed continuously from 0 to 10µm. It is important to know the linear relationship between the separation gap d and the position in the direction x (shown in figures above). The relationship was obtained using basic trigonometry. The major error source is the error of the width of the aluminium sheet:

d = Width of air gap

$$d = x\,(0{,}27 \pm 0{,}01)\,.\,10^{-3}$$

$$d(\mu m) = x(mm)\,(0{,}27 \pm 0{,}01)$$

Setup:



1. Laser (λ = 635 nm): The laser was placed so that light entered the first prism at an angle of incidence of 0^o, without changing direction of propagation. Consequently, the light encountered the hypotenuse of the first prism at a 45^o angle of incidence, which is slightly greater than the critical angle between the used glass and air ($41{,}3^o$).

2. Lens: One lens was placed in front of the laser to focalize the light, in order to make the beam thin enough to be considered punctual. Another lens was placed after the prisms for the same reason, avoiding the possibility of the beam becoming too large to be detected by the photocell.

3. Micrometer: All experimental setup was maintained still, except for the prisms. The prisms were placed on top of the micrometer, which made it possible to move them in the direction of x, causing light to reach the hypotenuse of the first prism at points of different position in x, and to measure this movement. Using the relationship between x and d, it is possible to calculate the width of the air gap at the point in which light is incident on the hypotenuse of the first prism.

4. Prisms: were put on top of the micrometer. The ones used (BK7) had a high optical quality.

5. Photocell: Used to capture the transmitted light.

6. Voltmeter: Attached to the photocell, this was used to measure the intensity of the transmitted light.

The prisms were moved in the x direction, and the intensity of the transmitted light beam was measured once each half millimetre moved. The experiment was realized in the dark, so that no external light influenced the measurements.

Results:

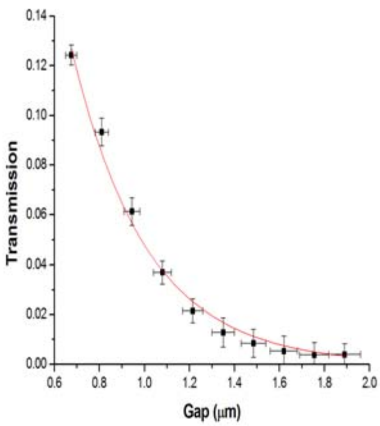
The table below shows the values obtained for each position in x:

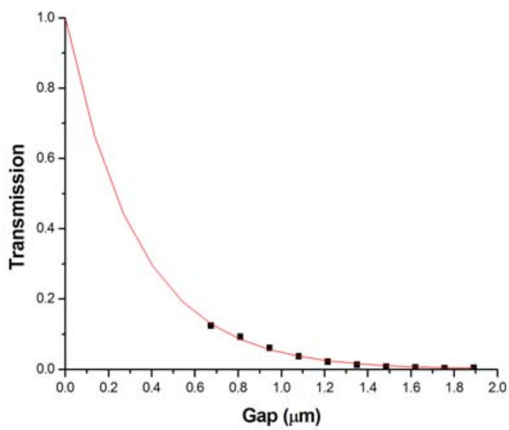
Position in X (mm)	Measure 1 (mV)	Measure 2 (mV)	Measure 3 (mV)	Average (mV)	Error (mV)
2,5	26,3	26,1	27,7	26,7	0,9
3,0	19,6	19,1	21,4	20,0	1,2
3,5	13,0	12,1	14,4	13,2	1,2
4,0	7,8	7,0	9,0	7,9	1,0
4,5	4,0	4,0	5,8	4,6	1,0
5,0	2,0	2,0	4,2	2,7	1,3
5,5	1,1	1,1	3,2	1,8	1,2
6,0	0,6	0,2	2,6	1,1	1,3
6,5	0,4	0,0	2,0	0,8	1,1
7,0	0,3	0,3	1,9	0,8	0,9

Measurements before 2,5 mm were not included due to their uncertainty, since the light incident at these regions was greatly scattered, so the intensity of transmission could not be measured in positions before 2,5 mm. With the data above, it was possible, using the program Origin 6.0, to obtain the equation which best describes the relationship of the voltage accused (I) as a function of the position in x. It is the equation which best adjusts to our experimental data and the error:

$$\mathbf{I = (210 \pm 30)e^{-x/(1,24 \pm 0,07)}}$$

where x is given in mm and I in mV. So, according to the equation, the voltage accused at x = 0 mm would be 210 ± 30 mV. Knowing this, the intensity of the transmitted beam was normalized so that this maximum intensity (at x = 0) would be equal to 1. It was possible to plot a graphic of the transmission (which can go from 0 to 1) versus the size of the gap:



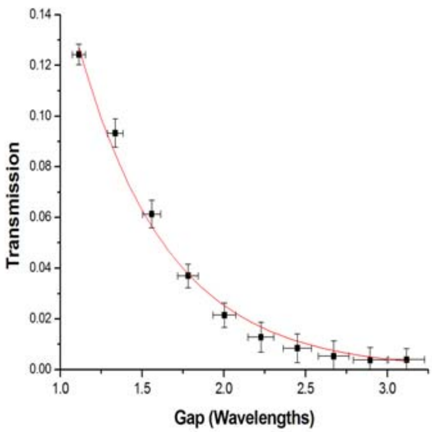


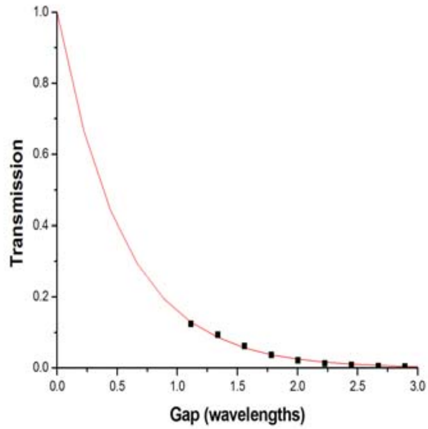
The dots represent the experimental values, and the line represents the exponential equation which best fits these results. The equation that best fits the experimental results is:

$$T = T_0 e^{-d/\,(0,33 \pm 0,02)}$$

where T stands for transmission and d is the gap in μm. The value **(0,33 ± 0,02) μm** would be our experimental value for 1/α (of the equation at the end of page 5), so the experimental value for α would be **(3,0 ± 0,2) μm^{-1}**. The expected value for α under the conditions of the experiment would be **3,79 μm^{-1}**, and 1/α would be **0,26 μm**. A possible explanation for the small difference between experimental and theoretical α would be that the width of the aluminium sheet would get smaller because it is being compressed by the two prisms. This is a non-quantifiable error source in the experiment. In fact, if we consider the width changed to 8 μm, we would have 1/α equal to (0,27 ± 0,2) μm accused by Origin.

Knowing the wavelength of the incident light used it was possible to build graphics with the gap distance in units of wavelength:





The last graphic proves the fading of the evanescent wave within a few wavelengths, which is predicted in theory. We can see that at a distance of about or greater than 2,5λ, the transmitted light is practically null.

Conclusions

- The phenomenon will occur if the distance between the prisms is of the order of the wavelength λ.
- Very well polished prisms are needed in order to perform the experiment with visible light.
- An application of the experiment would be the determination of a medium's index of refraction, because we can determine α.

References

- Castro, J.C. (1974) Optical Barrier Penetration: A Simple Experimental Arrangement. *American Journal of Physics*, **43**, 107-108.
- Coon, D.D. (1965) Counting Photons in the Optical Barrier Penetration Experiment, *American Journal of Physics*, **34**, 240-243.
- Mahan, A.I. and Bitterli, C.V. (1978) Total Internal Reflection: A deeper Look. *Optical Society of America*, **17**, 509-519.
- Hecht, E and Zajac, A. *Optics*. Addison-Wesley.
- Eisberg, R. *Mecânica Quântica*, Editora Campus.

- Newton, I. Óptica, EdUSP.
- OSA Unicamp *IYPT booklet*.

Acknowledgements

- Lino Misoguti (IFSC – USP)
- Ércio Santoni (IFSC – USP)
- Claudia E. Munte (IFSC – USP)
- Andreas Munte (Student at Educativa)
- Henrique Carvalho (UniCamp)
- Ducinei Garcia (UFSCar)
- Pedro Schio (UFSCar)
- Tatiane Godoy (UFSCar)
- Sérgio Motta (Teacher at Educativa)
- Ozimar da Silva Pereira (Chairman IYPT – Brazil)
- OSA Student Chapter Unicamp

APPENDIX

1. Resolution of electromagnetic wave equation.

The following equation comes from solving Maxwell's equations of electromagnetism. It describes the wave function of an electromagnetic radiation of frequency *f* propagating in a medium of index of refraction *n* in one dimension (x):

$$\frac{d^2\psi_{(x)}}{dx^2} + \left(\frac{2\pi f}{c} n\right)^2 \psi_{(x)} = 0$$

It can be rewritten as:

$$\frac{d^2\psi_{(x)}}{dx^2} + k^2\psi_{(x)} = 0 \qquad k = \frac{2\pi}{\lambda} = \frac{2\pi f}{c} n$$

In which k is called the *wave number*. Two possible solutions for this equation are:

$$\psi_{(x)} = \psi_{(0)} e^{\pm ikx}$$

It can be proved that these two solutions are possible. First, the proof for the positive exponential:

$$\frac{d^2\psi_{(x)}}{dx^2} = \psi_{(0)} i^2 k^2 e^{ikx} = -\psi_{(0)} k^2 e^{ikx}$$

$$k^2\psi_{(x)} = \psi_{(0)} k^2 e^{ikx}$$

Proceeding in the same manner for the negative exponential:

$$\frac{d^2\psi_{(x)}}{dx^2} = \psi_{(0)} i^2 k^2 e^{-ikx} = -\psi_{(0)} k^2 e^{-ikx}$$

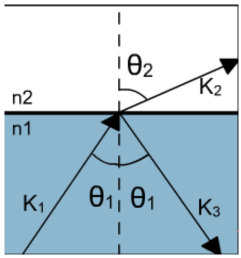
$$k^2\psi_{(x)} = \psi_{(0)} k^2 e^{-ikx}$$

A more general solution, for three dimensions, can be used to represent the spatial variation of an electromagnetic wave:

$$\psi = \psi_{(0)} e^{\pm i(\mathbf{k.r})}$$

where **k** is the wave vector or propagation vector, and **r** is the vectorial position.

2. Proof of Evanescent wave in total internal reflection.



Here we have an example of wave vectors **k** in refraction and reflection.

The equation describing the spatial variation of the electromagnetic wave is:

$$\psi = \psi_{(0)} e^{\pm i(\mathbf{k.r})}$$

The law of Snell can be applied whenever light encounters a boundary between two mediums:

$$\sin\theta_1 . n1 = \sin\theta_2 . n2$$

$$\sin\theta_2 = \frac{n1}{n2}\sin\theta_1 \longrightarrow \sin\theta_2 = \frac{\sin\theta_1}{\sin\theta_C}$$

$$\sin^2\theta_2 = \frac{\sin^2\theta_1}{\sin^2\theta_C} \longrightarrow 1 - \sin^2\theta_2 = 1 - \frac{\sin^2\theta_1}{\sin^2\theta_C}$$

$$\cos^2\theta_2 = 1 - \frac{\sin^2\theta_1}{\sin^2\theta_C} \longrightarrow \cos\theta_2 = \sqrt{1 - \frac{\sin^2\theta_1}{\sin^2\theta_C}}$$

Total internal reflection occurs when:

$$\theta_C < \theta_1 < 90^o \qquad \frac{\sin^2\theta_1}{\sin^2\theta_C} > 1$$

$$\cos\theta_2 = \sqrt{-1\left(\frac{\sin^2\theta_1}{\sin^2\theta_C} - 1\right)} \longrightarrow \cos\theta_2 = i\sqrt{\left(\frac{\sin^2\theta_1}{\sin^2\theta_C} - 1\right)}$$

$$\cos\theta_2 = i\sqrt{\left(\frac{n1}{n2}\right)^2 \sin^2\theta_1 - 1}$$

Having determined the cosine of θ_2 in total internal reflection, we may apply the wave equation to describe the wave that penetrates into the second medium in total internal reflection. To determine the behaviour of this wave in the direction d (normal to the surface) we must include in the wave equation a scalar product between **k** and the position in **d**:

$$\psi = \psi_{(0)} e^{\pm ikd(\cos\theta_2)}$$

$$\psi = \psi_{(0)} e^{\pm i(\mathbf{k}.\mathbf{d})}$$

$$ikd(\cos\theta_2) = d\, i^2 \frac{2\pi}{\lambda}\sqrt{\left(\frac{n1}{n2}\right)^2 \sin^2\theta_1 - 1} = -d\alpha$$

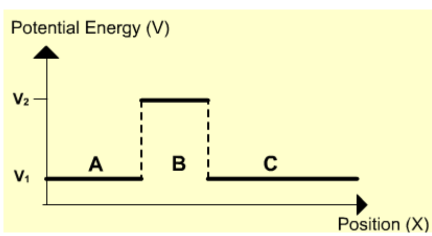
$$\alpha = \frac{2\pi}{\lambda}\sqrt{\left(\frac{n_1}{n_2}\right)^2 \sin^2\theta_1 - 1}$$

We have then 2 solutions, a positive and a negative exponential:

$$\psi = \psi_{(0)} e^{\pm d\alpha}$$

The positive exponential is discarded because it is physically impossible due to conservation of energy. The evanescent wave, therefore, decays exponentially.

3. Analogy to Quantum Physics: Every optical phenomenon has an analogue in quantum physics. Frustrated Total Internal Reflection is an optical analogue to particle potential barrier penetration (Quantum Tunnelling Effect). If a quantum particle (ex: an electron) encounters a potential barrier ahead of it during its movement, it will have a probability of being reflected from this barrier and a probability of passing through the barrier and therefore be encountered in the region beyond it.



Supposing our electron is moving in the sense of a growing position, in region A, and reaches a region of greater potential, B. We say it has encountered a potential barrier. It will have a probability either of returning back into region A and a probability of tunnelling into region C. This behaviour is different from macroscopic physics, in which we could predict if a body is or not to trespass a potential barrier. In quantum physics however, we must work with probabilities. The higher the potential barrier in quantum physics, the greater the probability of the particle being reflected, and the smaller the probability of tunnelling to occur. The length of the barrier also influences: the "longer" the potential barrier, the greater probability of reflection and the smaller the probability of tunnelling. In quantum mechanical optics analogies, the particles are the so – called photons that make up light. A beam of light can be well interpreted as a stream of photons. The quantum potential of a medium analogue to light will depend on the medium's index of refraction *n*.

$$\frac{d^2\psi_{(x)}}{dx^2} + \left(\frac{2\pi f}{c} n\right)^2 \psi_{(x)} = 0$$

This is the equation that describes electromagnetic radiation of frequency f propagating in a medium of index of refraction n. An equation applied in electromagnetism. Here ψ is the electric or magnetic field, x is position and *c* is the speed of light in vacuum.

$$\frac{d^2\psi_{(x)}}{dx^2} + \frac{2m}{\hbar^2}(E - V_{(x)})\,\psi_{(x)} = 0$$

This is Schrödinger's non time-dependent equation. It comes from quantum mechanics. Here ψ is the wave function, x is position, m is the mass of the photon (which is $h \,.f / c^2$, in which h is the Planck constant = $6{,}626 \,.\, 10^{-34}$ J.S), ħ is h/ 2π. E is the energy of the photon (h.f) and V is the quantum potential in the position x. The two equations express the same mathematical relation, so we can make an analogy between them. Comparing the equations, we obtain:

$$\left(\frac{2\pi f}{c} n\right)^2 = \frac{2m}{\hbar^2}(E - V_{(n)})$$

So we have the relationship between a medium's quantum potential analogue to light (V) as a function of it's index of refraction *n*:

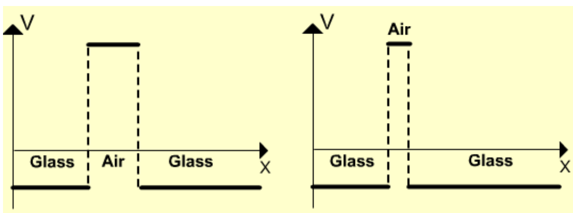
$$V_{(n)} = E - \frac{\left(\frac{fnh}{c}\right)^2}{2m}$$

By looking at the equation we can see that the bigger the medium's index of refraction *n*, the smaller the potential this medium offers to photons. Here we have some quantum potentials for red light (f = $4{,}48.10^{14}$ Hz):

Air (n = 1) = $1{,}45 \cdot 10^{-19}$ J
Water (n=1,3) = $3{,}21 \cdot 10^{-20}$ J
Glass (n = 1,5) = $-3{,}61 \cdot 10^{-20}$ J

These values depend on the frequency of the light considered. However, the **difference** between the potential offered by two mediums is the same for all frequencies.

The greatest potential is that of air. So we can say that the photon, when reaching the separation surface between glass and air in total internal reflection, encounters a potential barrier. If a second prism is placed at a distance of the order of λ, the length of the potential barrier will be small enough for quantum tunnelling of the photons to occur. The smaller the distance between the two prisms, the greater the probability of tunnelling of a single photon. So, the smaller the gap, the greater the intensity of light frustrated, because more photons tend to tunnel:



If the distance between the prisms is maintained, but the medium between them changes, the probability of tunnelling will also change because the height of the potential barrier will change. If we spread a fluid of greater index of refraction than air's, for example, water, on the surface of the prisms and put them macroscopically in contact, the distance between them will be the same as if there were air between them. However, the photons would encounter a potential barrier far smaller. Tunnelling would be made easier, because a same distance between prisms would offer a greater probability: