WAVE OPTICS

Jaan Kalda

Version: 12th April 2014

1 Basics. Double slit diffraction.

Visible light is an electromagnetic wave; in vacuum, its speed is constant and equal to $c = 3\times 10^8$ m/s; in a dielectric medium, the speed is reduced by a factor $n = \sqrt{\varepsilon}$, where $n = n(\omega)$ is the refraction coefficient, and ε is the relative dielectric permeability; both depend on the angular frequency of the electric field (here we assume that the magnetic permeability $\mu \approx 1$ for dielectric materials).

Maxwell equations admit several solutions; for instance, time-independent (stationary) solutions are possible. In particular, a point charge q creates an electrostatic field $\vec{E} = \frac{1}{4\pi\varepsilon_0}\frac{q}{r^2}\vec{e}_r$, where $\vec{e}_r$ is a unit vector pointing from the charge towards the observation point. Note that stationary electric fields are created by electric charges, and stationary magnetic fields — by electric currents. However, Maxwell equations include also terms with time derivatives (e.g. the time derivative of the magnetic flux in the Faraday's law); owing to these terms, wave-like solution are also possible. In particular, one can have a sinusoidal **plane wave**, for which the **wave fronts**[1] form a set of parallel planes:

$$\vec{E} = \vec{e}_x E_0 \cos(kz-\omega t),\ \ \vec{B} = \vec{e}_z B_0 \cos(kz-\omega t),$$

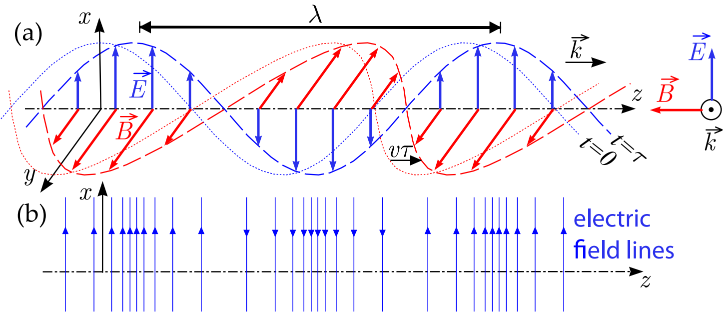
where z is the propagation direction axis,

$$k = 2\pi/\lambda, \tag{1}$$

is the wave vector, related to the circular frequency and wave speed via

$$\frac{\omega}{k} = v = \frac{c}{n} = \frac{c}{\sqrt{\varepsilon}}, \tag{2}$$

and the field amplitude ratio $E_0 = B_0 v$. Let us notice that at any point in space, the electric field, the magnetic field, and the propagation direction are all perpendicular to each other (propagation direction corresponds to the motion of a screw when rotated from E to B). In the figure (a) below, the vectors of the electric- and magnetic field are depicted for a certain time moment $t = \tau$, for a series of points lying on the z-axis; the endpoints of these vectors lay on sinusoids, which are drawn for $t = \tau$ and $t = 0$ (the dotted curves). In figure (b), electric field lines are depicted for the same wave.



In the complex number form, such a wave can be expressed as

$$\vec{E} = \vec{e}_x E_0 e^{i(\vec{k}\vec{r}-\omega t)},\ \ \vec{B} = \vec{e}_z B_0 e^{i(\vec{k}\vec{r}-\omega t)}, \tag{3}$$

where we have used the dot product of the radius vector $\vec{r} \equiv (x, y, z)$ with the wave vector $\vec{k}$ (which is parallel to the propagation direction of the wave). Alternatively, we can write

$$\vec{E} = \vec{e}_x E_0 e^{ik(z-vt)},\ \ \vec{B} = \vec{e}_z B_0 e^{ik(z-vt)}, \tag{4}$$

where we have substituted $\omega = kv$. Here, E_0 and B_0 can be complex numbers, so that $E_0 = |E_0|e^{i\varphi}$; then, E_0 is referred to as the complex amplitude, and $\varphi = \arg E_0$ is the wave's phase. Since the electric field of an electromagnetic wave defines immediately also the magnetic field, in what follows we consider only its electric field.

In real life, the wave fronts are not necessarily plane. In particular, a point source emits spherical waves, and a line source emits cylindrical waves. However, if the distance to the wave source of an arbitrary shape is much larger than the wave length, within a small neighbourhood of an observation point (of a radius of few wavelengths), the wavefront curvature is negligibly small[2]. Because of that, the interference of non-planar waves can be studied as the interference of locally plane waves. Still, an important aspect should be kept in mind: while for plane waves, the oscillation amplitude is constant throughout the space, for non-planar waves, the amplitude is a function of coordinates.

In particular, for spherical waves, the amplitude is inversely proportional to the distance from the point source, as it follows from the energy flux continuity. Indeed, the intensity of the wave (the energy flux density) is proportional to the squared wave amplitude, $I \propto E_0^2$; the energy flux (i.e. the total radiation power transmitted through a fictitious surface) equals to the product of the intensity, the surface area, and the cosine of the angle between the wave vector and the surface normal, $\Phi = IA\cos\phi$. With the origin being at the point source, let us consider the energy balance for the volume between two concentric spherical surfaces (of radii r_1 and r_2), within a solid angle Ω: the incoming energy flux equals to $\Omega r_1^2 I_1$, and the outgoing one — to $\Omega r_2^2 I_2$. In a stationary state and assuming that there is no energy loss due to dissipation, these two fluxes must be equal, i.e. $I_1 r_1^2 = I_2 r_2^2$ and hence, $E \propto \sqrt{I} \propto 1/r$. Similarly, for cylindrical wave, $E \propto 1/\sqrt{r}$.

As long as the propagation speed is constant, e.g. in a vacuum with $v = c$, any electromagnetic pulse will propagate with a constant shape and speed[3]:

$$\vec{E} = \vec{e}_x E(z-vt),\ \ \vec{B} = \vec{e}_z E(z-vt)/v$$

However, if n in Eq. (2) depends on the angular frequency ω, the pulse shape will change in time; furthermore, the speed of the pulse will be

$$v_g = \frac{d\omega}{dk} = \frac{d}{dk}\frac{ck}{n(k(\omega))}, \tag{5}$$

which is referred to as the group speed (then, ω/k yields the speed of a fixed phase, e.g. of a wave-crest, and is called the phase speed).

To understand why, let us consider the superposition of two beating waves of wave vectors $k-\Delta k$ and $k+\Delta k$, respectively:

$$e^{i[(k-\Delta k)z-(\omega-\Delta\omega)t]} + e^{i[(k+\Delta k)z-(\omega+\Delta\omega)t]} =$$
$$e^{i(kz-\omega)t}\left[e^{-i(\Delta kz-\Delta\omega)t} + e^{i(\Delta kz-\Delta\omega)t}\right] =$$
$$2e^{i(kz-\omega)t}\cos(\Delta kz-\Delta\omega t).$$

Here, the first factor $e^{i(kz-\omega)t}$ corresponds to the wave itself, and the second factor $\cos(\Delta kz-\Delta\omega t)$ describes its envelope; the speed of the envelope $v_g = \Delta\omega/\Delta k$.

In what follows, we consider only **monochromatic waves**, i.e. sinusoidal waves of a fixed frequency ω. This is because

[1] A wave front is defined as the set of points of constant oscillation phase (for instance, wave crests).
[2] Unless a concave source shape leads to a wave focusing near the observation point.
[3] The proof is provided in Appendix 1 on pg. 11.

we shall study the interference of light waves, and typically, an interference pattern can be observed only for light beams originating from a single source (this will be discussed in more details below). Now, if all the waves originate from the same source, they must have also the same frequency[4]. Note that if the wave enters a refractive transparent medium, the wavelength may change, but the frequency remains constant.

fact 1: The frequency of a wave remains constant along its entire path if the wave speed v depends only on coordinates and not on time (for light: if $n = n(x, y, z)$ does not depend on time).

Indeed, the time required for a wave crest to travel from the source to a given destination point is defined by the integral $\int \frac{dl}{v(x,y,z)}$, taken over the wave trajectory, which remains constant in time; hence, the time delay between neighbouring wave crests at the destination remains equal to what it was at the source.

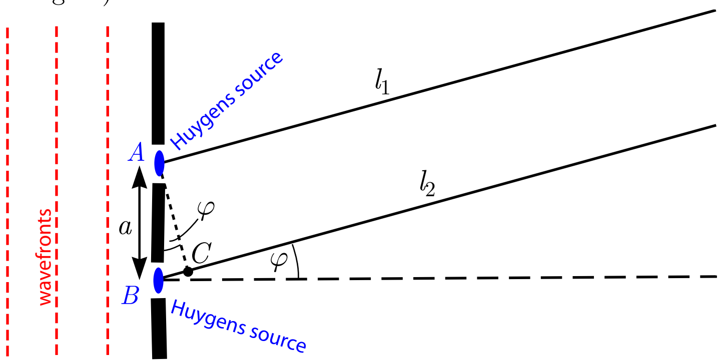
We also assume that the **coherence length** of the waves is larger than the system size. Coherence length is a distance upon which the wave "forgets" its phase. One can imagine this as having a sinusoid with slightly varying wavelength; upon certain distance, the variations accumulate into such an error that the phase difference between this wave and an ideal sinusoid will be of the order of π (which corresponds to an opposite phase). For the light sources other than lasers, the coherence length is really short; for lasers, it can reach the values around tens of meters.

fact 2: (The Huygens principle.) Consider an arbitrary wave propagation, for which a certain wave front is known. The wave propagates beyond that wave front in the same way it would propagate if a densely populated array of small wave sources were placed along the wave front.

This fact is the main tool for calculating diffraction patterns in majority of cases. Let us analyse this using the example of double slit diffraction (for some problems, including the first one, solutions are provided after the problem text).

pr 1. Consider a non-transparent wall with two parallel narrow slits (much narrower than the wavelength) at distance a from each other. Parallel light beam falls perpendicularly on that wall. Find the diffraction pattern behind the wall: the propagation angles for which there are light intensity maxima $\varphi_{\max}$ and minima $\varphi_{\min}$, as well as the intensity distribution as a function of the angle φ.

The wall blocks almost all the wave front of the original wave, leaving only two points in a cross-section perpendicular to the slits (see figure below). To be precise, these are actually segments, but their size is much smaller than the wavelength; so, from the point of view of wave propagation, the segments can be considered as points. According to the Huygens principle, two point sources of electromagnetic waves of wavelength λ will be positioned into these two points (A and B). The point sources radiate waves in all the directions, and we need to study the interference of this radiation. Let us study, what will be observed at a far-away screen where two parallel rays (drawn in figure) meet.



To begin with, it is quite easy to figure out, where are the intensity maxima and minima. Indeed, as it can be seen from the figure above, the optical path difference between the two rays is $\Delta l = a \sin\varphi$. The two rays add up constructively (giving rise to an intensity maximum) if the two waves arrive to the screen at the same phase, i.e. an integer number of wavelengths fits into the interval: $\Delta l = n\lambda$. Similarly, there is a minimum if the waves arrive in an opposite phase:

$$\sin\varphi_{\max} = n\lambda/a, \quad \sin\varphi_{\min} = (n + \frac{1}{2})\lambda/a. \tag{6}$$

Now, let us proceed with the calculation of the intensity distribution.

method 1: Quantitative calculation of the interference patterns is most conveniently done by adding up the complex amplitudes of interfering waves (similarly to alternating currents and voltages). Mathematically, if the wave amplitudes of M interfering are a_m, $m = 1, 2, \ldots M$, and the corresponding optical paths are denoted by l_m the resulting wave's complex amplitude is

$$a = \sum_m a_m e^{ikl_m}.$$

Note that complex numbers can be considered as two-dimensional vectors (the x- and y-coordinates of which are their real- and imaginary parts, respectively); because of that, alternatively, vector diagrams can be used (each wave is represented by a vector, the length of which reflects the wave amplitude, and the direction — the wave's phase). Here, the amplitudes a_m are proportional to the sizes of the Huygens sources. In the case of 3-dimensional geometry, they are and inversely proportional to the distance l_m, and in the case of 2-dimensional geometry — inversely proportional to $\sqrt{l_m}$.

To understand the origin of the source-size-proportionality, one can consider two identical near-by sources: due to the negligible distance, the corresponding waves have the same phase and therefore add up into a wave of double amplitude.

It should be noted that the formulation of this method ignores the dependence of the wave amplitude of the contributions of the Huygens sources on the propagation direction. This approximation is valid as long as the angle between the surface normal of a Huygens source and the vector pointing to the observation point is small (its cosine is approximately equal to one). If this angle is not small so that strictly speaking, dropping the angle-dependent factor would not be correct, doing so retains still the qualitative properties of the diffraction pattern if all

[4]The source can emit different frequencies, but such a radiation can be decomposed into a superposition of sinusoidal waves, as taught by the Fourier analysis.

the contributing beams are characterized by the same angle (like in the case of Pr. 1), because then, the angle-dependent factor has the same value for all the beams and can be brought before the braces.

In the case of light waves travelling in the z-direction, the amplitudes a_m and a are to be interpreted as x- or y-components of the E- or B-field. It is not important, which quantity is considered, because as long as there is no *double refraction*, for any contributing wave, there is no phase shift between E_x, E_y, B_x, and B_y. Meanwhile, for polarized light in double-refracting materials, x- and y-components need to be studied separately: then, the phase shift will depend on the oscillation axis; this will be discussed later. In order to avoid emphasizing which field is considered, in what follows the amplitudes will be denoted by a. Let us recall that the modulus of the complex amplitude gives the real amplitude of the sinusoidal signal, and the argument of it gives the phase shift. Thus, the real field $a(x,t)$, at the given point as a function of time, is given by

$$\mathrm{Re}\, a(l)e^{ikl} \cdot e^{-i\omega t} = |a(l)| \cos[\omega t + \arg a(l)].$$

Typically, momentary field values of electromagnetic waves are never needed: the oscillations are so fast that what is measured is the root-mean-square average. Therefore, the only things of interest are the modulus of the field, and its phase shift. Because of that, **it is enough to work with the complex amplitudes**; there is no need to write down the full spatio-temporal dependence of the complex signal $a(l)e^{i(kl-\omega t)}$, and there is no need to add Re to separate its real part which corresponds to a real physical quantity.

So, at our infinitely-remote-screen, we have two waves $a_1(l_1)e^{ikl_1}$ and $a_2(l_2)e^{ikl_2}$ adding up. The relative difference between l_1 and l_2 is small; hence, the dependence of the wave amplitude on distance affects the both waves in the same way, i.e. $|a_1(l_1)| = |a_2(l_2)|$. The two Huygens sources are at the same wavefront, which means that at the respective sources, there is no phase shift between the emitted waves, hence $\arg a_1(l_1) = \arg a_2(l_2)$; combining the last two equalities yields $a_1(l_1) = a_1(l_2)$. Since we are interested in the relative intensity of the light at the screen, and not in how it decreases with l, we can drop the dependence $a(l) \propto 1/\sqrt{l}$ and denote $a_1(l_1) = a_1(l_2) \equiv a$ (the sign "$\propto$" means "is proportional to"). Finally, we can combine the term e^{ikl_1} into the complex amplitude (this only rotates the complex ampitude as $|e^{ikl_1}| = 1$) by denoting $ae^{ikl_1} = \tilde{a}$, in which case

$$ae^{ikl_2} = \tilde{a}e^{-ikl_1} \cdot e^{ikl_2} = \tilde{a}e^{ik(l_2-l_1)}$$

We can also say that the amplitudes are normalized to the light wave amplitude from the first slit, and put $\tilde{a} = 1$[5].

So, the superposition amplitude is given by

$$E = 1 + e^{ik(l_2-l_1)}.$$

The intensity is proportional to the square of the modulus, which is given by the product of E with its complex conjugate $\bar{E} = 1 + e^{-ik(l_2-l_1)}$:

$$I/I_0 = (1 + e^{ik(l_2-l_1)})(1 + e^{-ik(l_2-l_1)}) =$$
$$= 2 + e^{ik(l_2-l_1)} + e^{-ik(l_2-l_1)} = 2\{1 + \cos[k(l_2 - l_1)]\};$$

here I_0 is such an intensity which would be recorded at the screan when one of the slits is closed. Alternatively, the square of the modulus can be calculated via Pythogorean theorem as the sum of the squares of the real- and imaginary parts:

$$\{1 + \cos[k(l_2 - l_1)]\}^2 + \sin^2[k(l_2 - l_1)] = 2\{1 + \cos[k(l_2 - l_1)]\}.$$

Finally, if we denote $l_2 - l_1 = a \sin\varphi$, we end up with

$$I = 2I_0[1 + \cos(ka \sin\varphi)]. \tag{7}$$

Now we can also recover the earlier result (6) regarding the positions of the intensity minima (cosine gives -1, $I = 0$) and maxima (cosine gives $+1$, intensity becomes quadruple): for minima, $ka \sin\varphi = (2n + 1)\pi$, and for maxima, $ka \sin\varphi = 2n\pi$.

Note that if there were non-coherent light sources at A and B, there would have been an additional time-dependent phase shift $\psi(t)$ which should have been added to the phase shift $ka \sin\varphi$ due to the optical path difference. In that case, $I = 2I_0\{1 + \cos[ka \sin\varphi + \psi(t)]\}$; owing to the fluctuating phase $\psi(t)$, the diffraction maxima (and minima) would move so fast that human eyes would register only the mean value of the intensity, $\langle 2I_0\{1 + \cos[ka \sin\varphi + \psi(t)]\}\rangle = 2I_0\{1+\langle\cos[ka \sin\varphi + \psi(t)]\rangle\} = 2I_0$; here angular braces denote averaging, and averaged cosine gives zero.

Regarding the fluctuations of the interference patterns from non-coherent light sources, we can make a simple estimation: let us have two point sources of fairly monochromatic light, for instance from two identical good lasers with a coherence length of $l = 10\,\mathrm{m}$ and wavelength $\lambda_1 = 658\,\mathrm{nm}$. Then the coherence time $\tau = l/c \approx 3 \times 10^{-8}\,\mathrm{s}$ gives us the fluctuation time of the random phase $\psi(t)$, and also the characteristic time-scale during which the diffraction pattern fluctuates. This is well beyond anything what an human eye can resolve: we'll see an averaged picture without any interference stripes.

So, as long as we are not studying phenomena at ultra-short time-scale (ato- and picosecond-scale-physics), in order to be able to see an interference pattern, **the light needs to come from the same light source**, even in the case of lasers. Additionally, the optical path differences of the interfering rays must not exceed the coherence length of the given light source. In the case of the double slit interference, it is sufficient if the light falls onto the slits from the same point source, not necessarily a laser. On the other hand, if it is not a point source, but instead a light bulb with a considerable size, the coherence length may become too short to be able to observe a two-slit-interference; a practical guideline here is that the angular size of the light source needs to be smaller than the angular distance between the diffraction maxima (otherwise two non-coherent halves of the source would give rise to two shifted diffraction patterns which become smoothed due to overlapping).

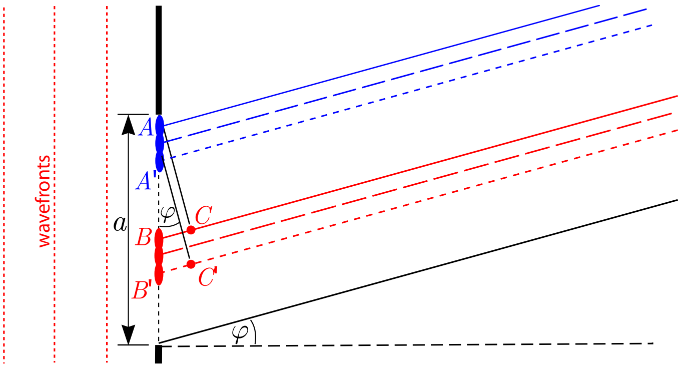
pr 2. Consider the same situation as in the case of pr 1, however with three slits at equal distance a from each other. Find the positions of diffraction minima and maxima.

pr 3. Show that for three parallel slits of equal size and with neighbour-to-neighbour distances being equal to a and b, respectively, the intensity at the diffraction minima is non-zero unless $\frac{a}{b} = \frac{n}{m}$, where n and m are integers and $n - m$ is a multiple of three.

[5] When solving wave optics problems, this paragraph can be summarized as on sentence: "at large distances, the contributing waves have equal amplitudes which will be normalized to the amplitude of a single wave

2 Single slit diffraction; diffraction grating

pr 4. Consider a non-transparent wall with a slit of width a. Parallel light beam falls perpendicularly onto that wall. Find the diffraction pattern behind the wall: the propagation angles for which there are light intensity maxima $\varphi_{\max}$ and minima $\varphi_{\min}$, as well as the intensity distribution as a function of the angle φ.



To begin with, let us find the positions of the diffraction minima — where the intensity is zero. To this end, let us divide the slit (fictitiously) into two halves; the Huygens sources from the upper half are marked with blue, and those from the lower half — with red, see figure. If the segment lenth BC equals to a half-integer-multiple of the wavelength then the contributions of the red and blue sources (from A and B) will cancel out; the same applies to any another matching pair, e.g. for points A' and B'. Indeed, the corresponding optical path difference $|B'C'| = |BC|$. So, the contributions from all the red and blue sources will cancel pair-wise out: the intensity is zero if

$$\frac{a}{2}\sin\varphi_{\min} = \lambda(n+\frac{1}{2}) \Rightarrow a\sin\varphi_{\min} = \lambda(2n+1).$$

Next, we divide the slit into four segments; then into eight segments, etc. in general, into 2^m segments; as a result, we find that a zero intensity is observed for

$$a\sin\varphi_{\min} = \lambda 2^{m-1}(2n+1).$$

One can see that here the factor of λ can take all the integer values except for zero[6]; so we can write

$$a\sin\varphi_{\min} = n\lambda \quad n \neq 0.$$

This result means that the main intensity maximum (at $\varphi = 0$) has a double width.

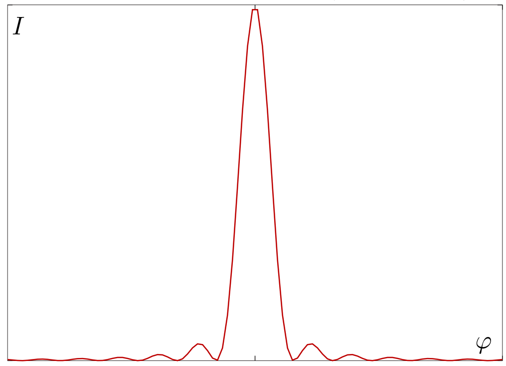
In order to find intensity distribution behind the slit, we need to integrate over the Huygens sources. Let us take the x-axis as the line AB, with the origin at B (i.e. at the centre of the slit). Then, each Huygens source contribution to the net wave amplitude E is proportional to its length dx; the optical path difference of this wave with respect to the wave arriving from B is given by $\Delta = x\sin\varphi$, which corresponds to the phase shift $x\sin\varphi$. Hence, the sum of all the waves can be written as

$$E \propto \int_{-\frac{a}{2}}^{\frac{a}{2}} e^{ikx\sin\varphi}dx = \frac{e^{i\frac{ak}{2}\sin\varphi} - e^{-i\frac{ak}{2}\sin\varphi}}{ik\sin\varphi} = \frac{2\sin(\frac{ak}{2}\sin\varphi)}{k\sin\varphi}.$$

Intensity is proportional to the squared amplitude, so that

$$I \propto \left[\frac{\sin(\frac{ak}{2}\sin\varphi)}{\sin\varphi}\right]^2$$

(we have dropped here constant factors 4 and k^{-2}). This dependence is shown in the figure below. Pay attention to the fact that dominating majority of the light energy is localized into the main maximum (between $a|\sin\varphi| < \lambda$).



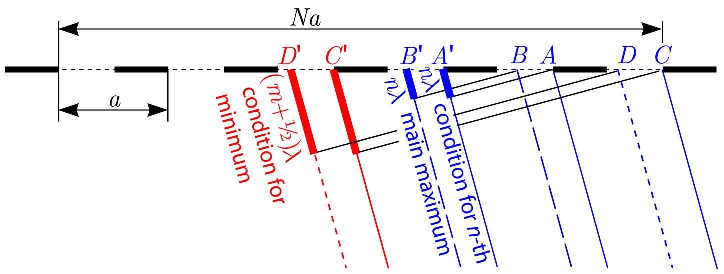
From the practical point of view, an important case is the diffraction behind a circular opening of diameter d — this will happen in the case of optical devices such as lenses, telescopes and microscopes. Finding the diffraction behind a circular opening of diameter d is mathematically significantly more challenging task, and involves Bessel functions; it appears that for $\lambda \ll d$, the first intensity minimum is observed for the angle $\varphi \approx 1.22\lambda/d$.

Let us assume that the front lens of a telescope (the *objective*) has a diameter d and creates an image of a twin star, the angular distance between the stars being α. It may be confusing that we have two effects: the diffraction on a circular opening, and ray convergence due to the lens. These two effects can be fortunately decoupled: suppose we remove the lens; then, the circular opening will create a diffraction pattern at an infinitely remote "screen". Now, we "put back" the lens, which creates the image of that infinitely remote pattern at its focal plane, according to the rules of geometrical optics.

It is said that the images of the two stars are resolved if the centre of the image of one star lies beyond the first intensity minimum of the diffraction pattern of the other star. According to angular position of the first diffraction minimum behind a circular hole, this means that the telescope resolves angular distances larger than $1.22\lambda/d$.

Next, let us consider a diffraction grating, which has N slits, neighbouring slits being at a distance a from each other.

pr 5. Calculate the diffraction pattern behind a diffraction grating assuming that the slit width is half of the grid pitch a.



Main maxima behind a grating can be found from the condition that the contributions from the neighbouring slits arrive at the same phase: the length of the thick blue lines in the figure needs to be an integer multiple of the wavelength, i.e.

$$a\sin\varphi_{\max} = n\lambda.$$

Apart from the main maxima, there are numerous side maxima; similarly to the case of a single-slit-diffraction, instead of finding the positions of these maxima, it is easier to find the positions of the minima; the maxima are just between the

[6] $m = 1$ gives all the odd numbers; $m = 2$ gives all such even numbers which give reminder 2 if divided by 4, etc

minima. The approach is also the same: we divide the whole diffraction grating into two halves, and consider the interference of the contributions of the both halves: the pair-wise cancellation of the light rays will take place if the length of the red thick line is an half-integer-multiple of the wavelength, i.e. $\frac{1}{2}aN\sin\varphi = (m+\frac{1}{2})\lambda$. Further we divide the grating into four, eight etc pieces, to conclude that the minima (with zero intensity) are observed for $aN\sin\varphi_{\min} = m\lambda$, where the integer $m \neq nN$ ($m = nN$ corresponds to the n-th main maximum).

In order to calculate intensity distribution behind such a grating, we can sum over the contributions of single slits. For the observation angle φ, we can use the expression for the electric field created by a single slit which was calculated earlier (we need to substitute a with $a/2$):

$$E_0 = \frac{\sin(\frac{ak}{4}\sin\varphi)}{\sin\varphi}.$$

Neighbouring slits have additional optical path difference $a\sin\varphi$, which corresponds to the phase shift $ka\sin\varphi$, and can be reflected by an additional term $e^{ika\sin\varphi}$ for the complex amplitude of the electric field. Thus,

$$E = \sum_{n=-N/2}^{N/2} \frac{\sin(\frac{ak}{4}\sin\varphi)}{\sin\varphi} e^{ikan\sin\varphi} =$$

$$= \frac{\sin(\frac{ak}{4}\sin\varphi)}{\sin\varphi} \sum_{n=-N/2}^{N/2} e^{ikan\sin\varphi}.$$

This is a geometric progression, and the sum can be easily taken:

$$E = \frac{\sin(\frac{ak}{4}\sin\varphi)}{\sin\varphi} \frac{e^{ika(\frac{N}{2}+1)\sin\varphi} - e^{-ika\frac{N}{2}\sin\varphi}}{e^{-ika\sin\varphi} - 1} =$$

$$= \frac{\sin(\frac{ak}{4}\sin\varphi)}{\sin\varphi} \cdot \frac{\sin\left(ka\frac{N+1}{2}\sin\varphi\right)}{\sin\left(\frac{ka}{2}\sin\varphi\right)} =$$

$$= \frac{\sin\left(ka\frac{N+1}{2}\sin\varphi\right)}{2\sin\varphi\cos\left(\frac{ka}{4}\sin\varphi\right)}.$$

Diffraction gratings are often used as spectral devices — to measure the spectrum of a light. In that case, it is important to have a good resolving power.

pr 6. Find the resolving power of a diffraction described by the previous problem, i.e. determine minimal value of $\Delta\lambda$, such that two spectral lines λ and $\lambda+\Delta\lambda$ can be resolved with such a grating.

In the case of a telescope, two points were assumed to be resolved, if the centre of one image lays beyond the first intensity minimum of the second image. In the case of a grating, we proceed in the same way: two spectral lines are resolved, if the centre of one line is beyond the nearest diffraction minimum of the other line. At the borderline case, these two things coincide; let the centre of one line's n-th main maximum be at φ; then

$$a\sin\varphi = n\lambda;$$

if this coincides with the nearest minimum of the second spectral line then

$$aN\sin\varphi = (nN-1)(\lambda+\Delta\lambda).$$

We can eliminate φ from these two equations to obtain

$$Nn\lambda = (nN-1)(\lambda+\Delta\lambda);$$

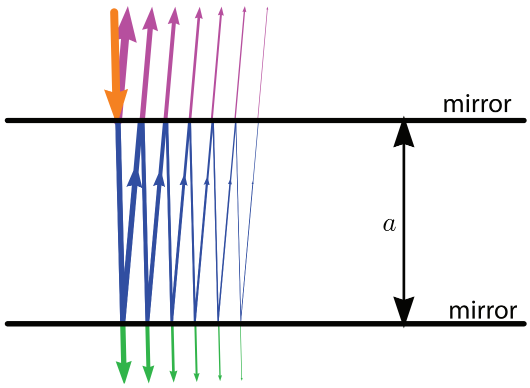
upon opening the braces and neglecting the term $\Delta\lambda$ for $N \gg 1$, we obtain

$$\frac{\lambda}{\Delta\lambda} = Nn.$$

Pay attention that at the right-hand-side of this expression, nN equals to the number of wavelengths which fit into the optical path difference between the shortest path and the longest path through our spectral device (for the direction at which the n-th main diffraction is observed). This is a very generic result, applicable to any spectral device, e.g. to a *Fabri-Perot'* or *Mach-Zehnder* interferometer or to an *echelle grating*.

The last result implies that larger physical size of a grating provides a better resolving power. However, in practice, this is not always the case. The reason is that the above derived formula assumes that the grating is ideal, with a strictly constant pitch. In practice, however, the pitch can fluctuate, and in that case the limiting factor will be the coherence length of the stripes — the length l which can be covered either by N or $N+1$ stripes, the uncertainty being due to the pitch fluctuations.

pr 7. Estimate the resolving power of a Fabri-Perot interferometer as a spectral filter, and find its spectral transmittance as a function of wavelength. This interferometer consists of two semi-transparent and semi-reflecting parallel surfaces with a very high reflectivity R (this gives the fraction of light energy which is reflected — as compared with the incident beam's energy), which are separated by a distance a.



The resolving power can be estimated easily using the above-mentioned generic rule. The shortest optical path is the one which goes through the interferometer directly without being reflected; the longest one performs many reflections. Strictly speaking, such a multiply-reflected beam has no upper length limit, but too many reflections lead to an almost vanishing intensity of the beam. A beam will participate efficiently in diffraction, if its intensity is not much smaller than that of the other beams. The number of reflections can be estimated as $N \approx \frac{1}{1-R} \gg 1$. Indeed, upon N reflections, the remaining intensity of the light is reduced by a factor of R^N; let us take the bordeline value for this factor to be e^{-1}. Then $R^N \approx e^{-1}$ from where $N \approx -\frac{1}{\ln R} \approx \frac{1}{1-R}$. So, we obtain

$$\frac{\lambda}{\Delta\lambda} \approx \frac{1}{1-R}.$$

The spectral transmittance can be found in two ways. The first way is to sum the contributions of several reflections using the formula for a geometric progression. The second way is to combine all the upwards propagating waves into a single wave, and all the downwards propagating waves into another wave. Then, the amplitudes of these effective waves can be tailored to each other via the reflection condition (see below), and eventually expressed in terms of the incident wave amplitude. This approach is valid because of the following idea.

idea 1: The sum of several sinusoidal waves of equal wavelength propagating in the same direction is also a sinusoidal wave propagating in the same direction.

Indeed, let us have N sinusoidal waves, and let the n-th wave be represented in a complex form as

$$a_n(z) = A_n e^{i(kz-\omega t)},$$

where A_n its complex amplitude. Then, the sum of the waves is given by

$$a(x) = \sum_n A_n e^{i(kz-\omega t)}.$$

The exponential term is the same for all the waves and hence, can be brought before the braces (i.e. the summation sign):

$$a(x) = e^{i(kz-\omega t)} \sum_n A_n.$$

Now, $\sum_n A_n$ is a complex number, let us denote it by A. Then, $a(z) = Ae^{i(kz-\omega t)}$, i.e. we have a wave with the same wavelength and the same direction of propagation as the component-waves.

Thus, we can combine all the purple up-moving-waves (see the figure above) into a single wave of amplitude E_r, and all the green down-moving-waves into a single wave E_t (indices t and r standing for "transmitted" and "reflected"). Similarly we can combine all the blue up-moving waves into E_u and all the blue down-moving-waves into E_d. Finally, let the amplitude of the incident wave (the red one) be denoted by E_i. Then we can say that E_t is the transmitted part of E_d, and E_u is the reflected part of E_d:

$$E_t = \sqrt{1-R}E_d, \quad E_u = \sqrt{R}E_d.$$

Note that we need a square root because reflectivity R and transmittance $1-R$ are related to the intensity (we deal with the amplitudes, which are proportional to the square root of the intensity). Further, E_d is made up from the transmitted part of E_i and reflected part of E_u. Now we need to pay attention to the phase shift between these three complex amplitudes. The phase shift of a wave will be changed by kl, if we shift the origin by distance l; since the "blue down wave" and the "orange incident wave" propagate in different (non-overlaping) regions , we can use different origins for them and make the phase shift of both waves equal to zero. Thus, we can assume that the transmitted component of E_i contributes to the complex amplitude of the "blue down wave" without phase shift[7], with $\sqrt{1-R}E_i$. Then, upon travelling down and up, the optical path length accumulated by the blue wave equals to $2a$, and hence, the reflected component of E_u comes with a phase factor e^{2ika}:

$$E_d = \sqrt{1-R}E_i + \sqrt{R}E_u e^{2ika}.$$

Using these three equations, we can express E_t in terms of E_i. Indeed, $E_d = E_t/\sqrt{1-R}$ and $E_u = E_t\frac{\sqrt{R}}{\sqrt{1-R}}$; thus,

$$\frac{E_t}{\sqrt{1-R}} = \sqrt{1-R}E_i + \frac{R}{\sqrt{1-R}}E_t e^{2ika},$$

from where

$$E_t\left(1-Re^{2ika}\right) = (1-R)E_i \Rightarrow E_t = \frac{1-R}{1-Re^{2ika}}E_i.$$

By definition, effective transmittance $t = |E_t|^2/|E_i|^2 = E_t\bar{E}_t/E_i\bar{E}_i$; therefore,

$$t = \frac{1-R}{1-Re^{2ika}} \cdot \frac{1-R}{1-Re^{-2ika}} = \frac{(1-R)^2}{1+R^2-2R\cos(2ka)}.$$

Let us pay attention to the fact that at the transmission maximum, $t = 1$, and at the transmission minimum, $t = \left(\frac{1-R}{1+R}\right)^2$.

The reflected wave (the purple up-moving one) can be found from the energy conservation law: the effective reflectivity $r = 1-t$. Alternatively, it can be found as the superposition of the reflected part of E_i and the transmitted part of E_u. Here, however, we need to take into account additional phase shifts during reflections.

fact 3: If electromagnetic wave is reflected from the interface of two dielectric media, it will partially reflected and partially refracted (as long as it is not a total internal reflection); in optically sparser medium, the reflected wave obtains an additional phase shift of π at the interface. There is no phase shift for other waves (refracted waves and reflected beam in the optically denser medium).

Partial reflection can take place on various interfaces, for instance on very thin metal films. It can be proved using the energy conservation law that regardless of what kind of interface there is, the sum of phase shifts between the transmitted and reflected waves (for the both directions of the incident wave) equals to π. In the case of thin metal film, there is a mirror symmetry, so the phase shifts cannot depend on the direction of incidence and hence, the phase shift between the transmitted and reflected waves is $\pi/2$.

pr 8. Prove that for an arbitrary semi-reflecting dissipationless interface, the sum of phase shifts between the transmitted and reflected waves for the both directions of the incident wave equals to π.

Returning to the case of Fabri-Perot interferometer, one can assume that the mirrors are dielectric and between the mirrors, there is optically denser medium. Then, there is a phase shift π between the red and purple waves (see figure above), so that $E_r = \sqrt{1-R}E_u e^{2ika} - \sqrt{R}E_i$.

3 Bragg reflection

The interaction of X-rays with ordinary matter is typically very weak. This is because the frequency of X-rays is much higher than the natural frequencies of electrons around the atoms and molecules. What happens is completely analogous to the mechanical resonance which is described by its equation of motion

$$m\ddot{x} = -kx + f_0\cos(\omega t),$$

where m is the mass of a particle attatched to a spring of stiffness k, x is the displacement of the particle, and $f_0 e^{i\omega t}$ — the external forcing. If the forcing frequency is close to the natural frequency of an oscillator, the oscillation amplitude will become very large; for low frequencies, the oscillator will take a quasi-equilibrium position: the displacement of the system is given by the momentary value of the forcing, $x = f_0\cos(\omega t)/k$. For very high frequencies, the strain force of the spring will play a negligible role and the oscillator will behave as almost a free particle, $m\ddot{x} = f_0\cos(\omega t)$, which can be integrated twice to yield $x = -f_0\cos(\omega t)/\omega^2$. This result means that the oscillation amplitude will decrease inversely proportionally to the squared

[7] The semitransparent mirror can cause a phase shift, but in that case, we use an appropriately shifted coordinate systems for the interior region.

frequency, and the displacement vector of the oscillator will be in the opposite phase with the forcing.

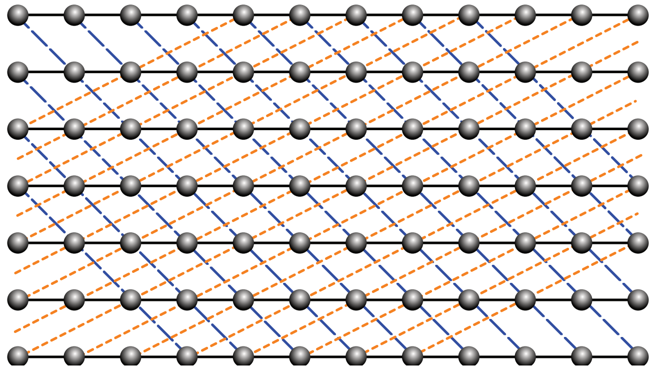
In the case of low-frequency electromagnetic waves (such as radio waves), the forcing frequency is much smaller than the natural frequencies of the electrons, the molecules will be deformed and polarized exactly in the same way as when being put into an electrostatic fields. Hence, the frequency-dependent dielectric permeability takes its stationary value. In particular, the refractive index of water for such waves is $n = \sqrt{\varepsilon(0)} \approx 9$.

Near natural frequencies (close to a resonance), the wave energy is pumped into oscillations of electrons, which leads to dissipation: matter becomes opaque. For frequencies much higher than the natural frequencies of the orbital electrons (e.g. for X-rays), the electrons respond to the forcing in the same way as free electrons (not bound to molecules). Indeed, the natural frequencies are of the same order of magnitude as the orbital rotation frequency, which means that during one forcing period, the orbital displacement of electrons remains much smaller than the orbital radius. We have seen above that for a free particle in a sinusoidal force field, the displacement is in the opposite phase with the forcing. In the case of low frequencies, the displacement is in the same phase with the forcing, which leads to a decrease of the overall electric field due to the polarization of the molecules; this effect is described by the relative dielectric permeability $\varepsilon > 1$. In the case of high frequencies, the effect is opposite, hence $\varepsilon(\omega) < 1$, and also $n(\omega) = \sqrt{(\varepsilon)} < 1$. Such values imply that the phase speed of light $v = c/n > c$, which may seem to be in a contradiction with the theory of relativity; however, relativistic constraint applies to the energy and information transfer rate only, which is given by the *group speed* of electromagnetic waves.

Unlike in the case of gamma rays, the frequency of X-rays is lower than the natural frequencies of nuclear oscillation modes, hence they interact only with the orbital electrons (to be correct, they do interact with the whole nuclei as with charged particles, but the mass of a nucleus is much larger than the electron mass, so the interaction is much more efficient in the case of electrons). As we have argued, the electrons behave as free particles, hence the interaction strength is defined only by the volume density of electrons: higher density of electrons implies larger value of $n-1$. In particular, this is why iridium is used for the mirrors of X-ray telescopes (even though $n-1$ remains still small, for small grazing angles, total internal reflection can be achieved).

A weak refraction is not the only way in which the X-rays can interact with matter. There is also a possibility that they are absorbed or scattered from the electrons; in that case, an X-ray beam behaves as a beam of particles — photons, which collide with the electrons (in the case of absorption, an electron at a lower orbital receives the photon's energy, and "jumps" to a free orbital). These are probabilistic effects; the scattering (or absorption) probabilities are to be calculated using equations of quantum mechanics. Scattering of photons on electrons (when electrons and photons are considered as elastic balls) is studied in the section "Quantum mechanics"; calculation of scattering probabilities is beyond the scope of the IPhO Syllabus.

Finally, X-rays can be reflected by a regularly arranged array of ions (present in crystals); this phenomenon called **the Bragg reflection**. According to the Bragg model, **crystal planes** perform as weakly reflecting (mostly transparent) surfaces. Crystal planes are fictitious planes at which a large number of ions is situated (see figure with a cubic lattice: cross-sections of crystal planes are depicted by black solid, blue dashed and orange dotted lines); the planes with higher ion surface density reflect X-rays more efficiently (in the figure, surfaces marked with black solid lines have the highest density).

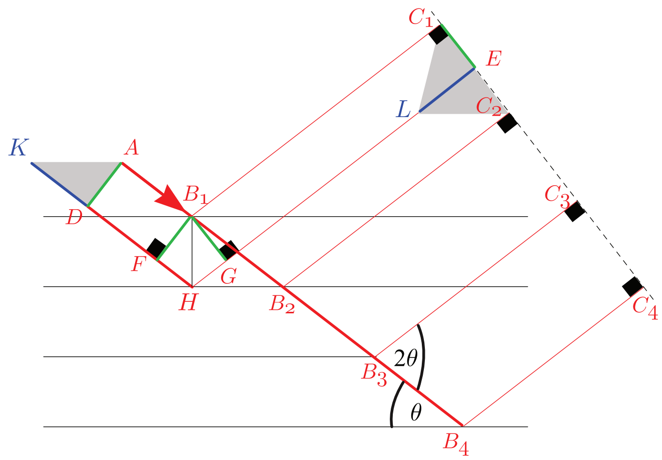


Although each surface reflects only a small amount of light, if the reflections from many surfaces add up in the same phase, the total reflected light can be significant; in fact, almost all the light can be reflected. So, the condition of Bragg reflection is that the contributions from neighbouring crystal surfaces are in the same phase, i.e.

$$\Delta l = |AB_2| + |B_2C_2| - |AB_1| + |B_1C_1| = n\lambda.$$

Calculating Δl in such a way is not easy enough — it can be done in an easier way. Indeed, instead of studying the reflections of the same ray AB_1, let us consider different rays of the incident beam of X-rays, AB_1 and KH.

idea 2: Different rays of the incident beam have the same phase at those points which lay on the same wave front. Wave front is a surface perpendicular to the rays.



Therefore, the incident wave has the same phase at the points A and D; the same applies to the pair of points B_1 and F. Similarly, for the reflected wave, C_1 and E have the same phase, and B_1 and G have the same phase. So, Bragg reflection condition can be written as condition that

$$\Delta l = |FH| + |HG| = 2a \sin\theta = n\lambda,$$

where a is the distance between neighbouring crystal planes. This is the main formula for Bragg condition: the X-rays are reflected by a crystal if the grazing angle θ between the X-rays and the crystal plane satisfies the condition

$$\sin\theta = \frac{n\lambda}{2a}.$$

Pay attention to the fact that for the same crystal, a can take different values and can be arbitrarily small (e.g. in the figure above, purple planes are significantly closer than the red ones). However, there is always a maximal value for a (corresponds to the red planes above). So, a crystal can reflect X-rays under many angles, due to the different values of a and n; the smallest value of θ corresponds to $n = 1$ and to the largest value of a. Finally, don't forget that the angle by which the X-rays are deflected equals to 2θ.

Note that the equality $|AB_2| + |B_2C_2| - |AB_1| - |B_1C_1| = |FH| + |HG|$ can be also verified geometrically. Indeed, $|KH| + |HL| = |AB_2| + |B_2C_2|$; due to the congruence of grey triangles, $|KH| + |HL| = |DH| + |HE|$. Thus, $|AB_2| + |B_2C_2| = |DH| + |HE|$, i.e. $|AB_2| + |B_2C_2| - |AB_1| - |B_1C_1| = |DH| + |HE| - |AB_1| - |B_1C_1| = |DH| + |HE| - |DF| - |GE| = |FH| + |HG|$.

4 Polarization. Double refraction.

Up till now we have implicitly assumed the light to be linearly polarized — by assuming a fixed axis (x) for the direction of electric field. Natural light, however, is in most cases non-polarized. This means that the direction of the electric field fluctuates in time. This is effectively another aspect of non-coherence: after a certain time period (it may be many wave periods, but in seconds, still a really tiny amount). The electric field of the electromagnetic wave "forgets" its previous direction and takes a new arbitrary direction (perpendicular to the direction of propagation, see Section 1). This happens so fast that neither human eye nor common measuring devices are able to discern the momentary directions of the electric field. From the point of view of diffraction studied above, this is not really important since two coherent light beams (from the same source!) have the same momentary direction of the electric field (as long as the optical path difference of the beams does not exceed the coherence length). So, non-polarized light is an electromagnetic wave which has randomly fluctuating direction of the electric field.

However, it is possible to have light, the electric field of which is always parallel to a fixed axis; the plane defined by this axis and propagation direction is called the polarization plane. In particular, when light is reflected by a dielectric interface under *Brewster angle* $\alpha_B = \arctan n$, the reflected beam is completely polarized (electric field is parallel to the interface). In this case, the reflected an refracted beams are perpendicular to each other. The refracted beam is also polarized, but only partially (most of the light is polarized perpendicularly to the interface). Partially polarized light can be thought as a superposition of two non-coherent polarized waves with perpendicular planes of polarization.

Light reflected by a dielectric surface is always somewhat polarized, and the non-polarized component decreases as the incidence angle approaches the Brewster angle. Similarly, the blue light from the sky is also partially polarized. This is because we see blue light from the sky due to Rayleigh scattering. If we consider a small fictitious volume of air, the number of molecules in it fluctuates somewhat; if the average (expected) number of molecules is N, typical fluctuations in the number of molecules are around $\sqrt{N}$. So, the relative fluctuations in the density of air are of the order of $1/\sqrt{N}$: they grow with decreasing N. The departure of the air's coefficient of refraction n from unity is proportional to the density of air. Hence, small fictitious volumes of air behave as media of different coefficient of refraction: there is partial reflection from these volume boundaries. The amount of reflected light is still small, because the difference in n is small. The effect is stronger for smaller fictitious volumes, but a wave cannot "discern" anything smaller than *ca* a quarter of the wavelength. This is the reason why the sky is blue: blue light has shorter wavelength and hence, can "see" smaller volumes with higher fluctuations in n than the other components of the sunlight. Now, what we see as a blue sky is a light reflected by a dielectric "interface", which is partially polarized. The polarization is the strongest for Brewster angle, when the reflected and refracted beams are perpendicular. Since n is very close to one, the refracted beam goes almost along a straight line, parallel to a vector pointing to Sun. So, in the case of the Brewster angle, the reflected (scattered) light is perpendicular to the direction of Sun: if you look into sky perpendicularly to the Sun, you see a strongly polarized blue light.

There are materials which have the so called double refraction property; for a linearly polarized light, the coefficient of refraction depends on the polarization plane. Furthermore, some materials are transparent for one polarization plane, and opaque for the perpendicular one. These materials are used to make linear polarizers — filters which let through only a light which is polarized in a specific plane. When a non-polarized light goes through such a filter, half of the light energy is dissipated (the light which was polarized in a wrong direction), and at the output, we have a completely polarized light. Such filters are used in photography to reduce reflections from dielectric surfaces (such as water or glass); in the case of Brewster angle, the reflections can be removed entirely. Also, these filters can make sky darker and remove blue haze obscuring distant objects (e.g. mountains); don't forget that such a haze-removal works best if the Sun is perpendicular to the direction of observation (see above).

The Brewster reflection can be used for precise measurements of the coefficient of refraction. For instance, when a totally polarized laser light falls onto a dielectric surface, the reflected beam disappears for Brewster angle α_B assuming that the polarization plane of the laser light is perpendicular to the surface. Then, α_B can be measured, and n is found as $n = \tan \alpha_B$.

Now, consider a case when a polarized light falls onto a polarizer so that the polarization planes form an angle α. Let the polarizer's polarization plane define the x-axis; suppose that before the polarizer, the electric field vector at its maximum is $\vec{E}_0$. This vector can be represented as a sum of two vectors $\vec{E}_0 = \vec{e}_x E_0 \cos\alpha + \vec{e}_y E_0 \sin\alpha$, which represents the decomposition of the initial wave into two perpendicularly polarized components. The polarizer dissipates completely the y-component, and at the output we have electric field amplitude vector equal to $\vec{e}_x E_0 \cos\alpha$ ($\vec{e}_x$ and $\vec{e}_y$ are the unit vectors along x and y axis, respectively). Let us recall that the intensity is proportional to the squared amplitude; therefore, the transmitted light's intensity

$$I = I_0 \cos^2 \alpha,$$

which is referred to as the Malus' law (I_0 is the intensity of the incoming polarized light).

pr 9. Let us have two polarizers with perpendicular planes of polarization. A non-polarized light beam of intensity I_0 falls onto such a system, and of course, no light can be detected at the output. Now, a third polarizer is inserted between the two polarizers, so that it forms an angle α with the polarization plane of the first polarizer. What is the intensity of light at the output?

Apart from the non-polarized light and linearly polarized light, there is also a circularly polarized light and elliptically polarized light. These can be obtained from the linearly polarized light by using double refracting plates. As mentioned above, the refraction coefficient of double refracting materials depends on the polarization plane. In the case of a so-called quarter-wavelength plate, this effect leads to the optical path difference between two components equal to $\frac{\lambda}{4}$. Let us choose x and y axis at the plate's plane so that the y-polarized light is retarded with respect to the x-polarized one by $\frac{\lambda}{4}$. Let us have linearly polarized light falling onto such a plate so that the polarization plane forms an angle α with the x-axis. Then, at a certain point before the plate, the time-dependence of the electric field components are given by

$$E_x = E_0 \cos\alpha \sin(\omega t), \quad E_y = E_0 \sin\alpha \sin(\omega t).$$

Thus, at any moment of time, $E_y/E_x = \tan\alpha$, i.e. the electric field vector oscillates along a line $E_y = E_x \tan\alpha$. After the plate

$$E_x = E_0 \cos\alpha \sin(\varphi_0 + \omega t + \frac{\pi}{2}), \quad E_y = E_0 \sin\alpha \sin(\varphi_0 + \omega t),$$

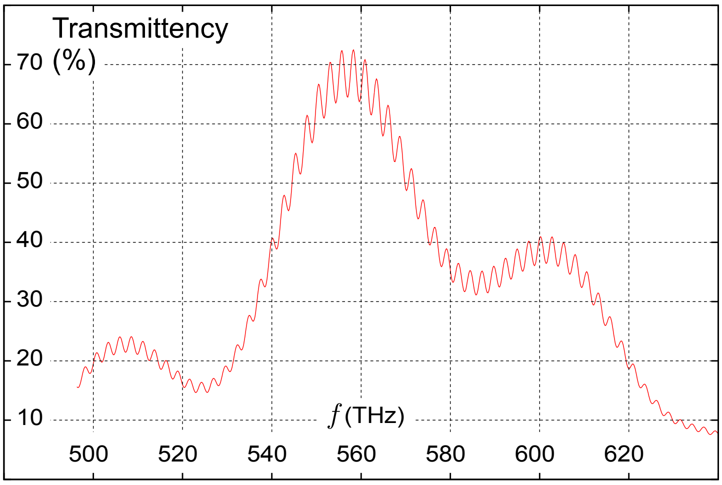
where φ_0 is the phase depending on the point of observation. From this system of equations, we can easily obtain

$$\left(\frac{E_x}{\cos\alpha}\right)^2 + \left(\frac{E_y}{\sin\alpha}\right)^2 = E_0^2,$$

which is the equation of an ellipse: the endpoint of the electric field vector draws such an ellipse. Therefore, such a light is said to be elliptically polarized; in the particular case of $\alpha = \frac{\pi}{4}$, it is circularly polarized.

Sometimes it is important to use a linear polarizer, but to avoid linearly polarized light at the output (e.g. in modern cameras semi-transparent mirrors are used to split the light, and if the light is linearly polarized, the balance between the intensities of split beams becomes non-predictable). Then, a quarter-wavelength plate is attached to a polarizer so that the output light becomes circularly polarized; these are called *circular polarizers*.

pr 10. A thick glass plate is coated by a thin transparent film. The transmission spectrum of the system is depicted in graph (light falls normal to the plate). The refractive index of the film $n \approx 1.3$. What is the thickness of the film d?

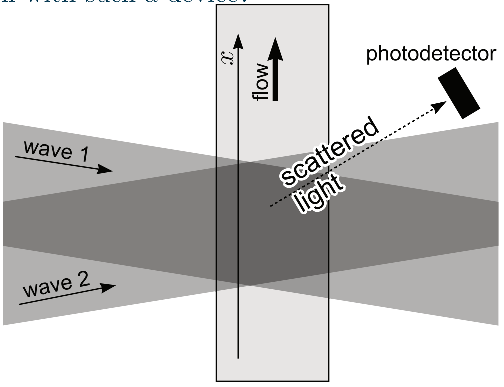


pr 11. Anemometer is a device measuring flow rate of a gas or a fluid. Let us look the construction of a simple laser-anemometer. In a rectangular pipe with thin glass walls flows a fluid (refractive index $n = 1{,}3$), which contains light dissipating particles. Two coherent plane waves with wavelength $\lambda = 515\,\text{nm}$ and angle $\alpha = 4°$ between their wave vectors, are incident on a plate so that *(a)* angle bisector of the angle between wave vectors is normal to one wall of the pipe and *(b)* pipe is parallel to the plane defined by wave vectors. Behind the pipe is a photodetector, that measures the frequency of changes in dissipated light intensity.

bf (i) How long is the (spatial) period Δ of the interference pattern created along x-axis (see Figure)?

(ii) Let the oscillation frequency of the photometer signal be $\nu = 50\,\text{kHz}$. How large is the fluid's speed v? What can be said about the direction of the fluid flow?

(iii) Let us consider a situation, when the wavelengths of the plane waves differ by $\delta\lambda = 4{,}4\,\text{fm}$ (1 fm$= 10^{-15}$ m). What is the frequency of signal oscillations now (fluid's speed is the same as in previous section)? Is it possible to determine the flow direction with such a device?



pr 12.

As it is well known, a telescope makes it possible to see the stars in daylight. Let us study the problem in more details. Consider a simplified model of the eye: a single lens with focal length $f = 4\,\text{cm}$ and diameter $d = 3\,\text{mm}$ creating an image on screen (retina). The model of a telescope is similar: a lens of focal length $F = 2\,\text{m}$ and diameter $D = 20\,\text{cm}$ creating an image in focal plane (where eg. a film can be put). In your calculations, the following quantities can be used: the density of the light energy radiated from a unit Solar surface in unit time w_0 (the light power surface density); the ratio of the star and Sun distances $q = 4 \cdot 10^5$ (we assume that the star is identical to the Sun); Solar angular diameter $\phi \approx 9\,\mathrm{mrad}$. *Remark:* If the answer contains w_0 then numerical answer is not required.

(i) Consider a sheet of paper, the normal of which is directed towards the Sun. What is the surface density of the light power w_1 arriving to the sheet from the Sun?

(ii) Find the net power P_2 of the light, which is focused by the telescope into the image of the star.

(iii) Assume that blue sky is as bright as a sheet of grey paper illuminated by Sun. You may assume that in the direction, perpendicular to the sheet, the ratio of the light power scattered by the paper into a 1-steradian space angle, to the net light power arriving to the sheet, is $\alpha \approx 0{,}1$ (this corresponds to the dissipation of ca 70 % light energy in the grey paper). What is the surface density of the light power in the focal plane of the telescope w_3, due to the blue sky?

(iv) While studying the star image, let us ignore all the effects other than diffraction. Estimate the surface density of the light power in the centre of the star image w_2 (in the focal plane of the telescope), due to the light arriving from the star.

(v) Provide an expression for the ratio of the surface densities of the light powers k in the middle of the star image, and in a point farther away from it.

(vi) Is it possible to see a star in daylight using a telescope? Plain eye? Motivate yourself.

pr 13. [IPhO-1981] A detector of radiowaves in a radioastronomical observatory is placed on the sea beach at height $h = 2\,\mathrm{m}$ above the sea level. After the rise of a star, radiating electromagnetic waves of wavelength $\lambda = 21\,\mathrm{cm}$, above the horizon the detector registers series of alternating maxima and minima. The registered signal is proportional to the intensity of the detected waves. The detector registers waves with electric vector, vibrating in a direction parallel to the sea surface.

(i) Determine the angle between the star and the horizon in the moment when the detector registers maxima and minima (in general form).

(ii) Does the signal decrease or increase just after the rise of the star?

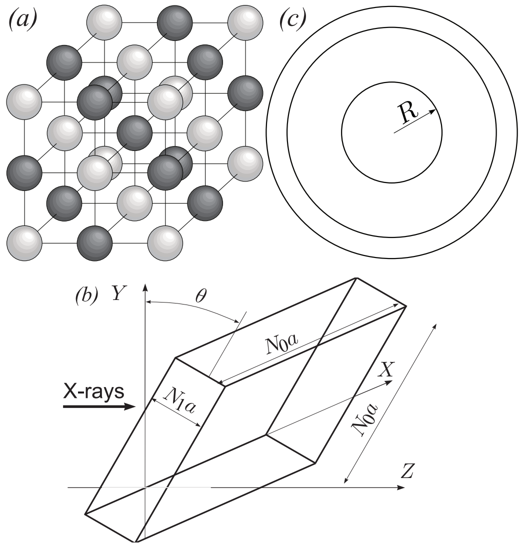
(iii) Determine the signal ratio of the first maximum to the next minimum. At reflection of the electromagnetic wave on the water surface, the ratio of the intensities of the electric field of the reflected (E_r) and incident (E_i) wave follows the low:

$$E_r/E_i = (n - \cos\varphi)/(n + \cos\varphi),$$

where n is the refraction index and φ is the incident angle of the wave. For the surface "air-water" for $\lambda = 21\,\mathrm{cm}$, the refraction index $n = 9$.

(iv) Does the ratio of the intensities of consecutive maxima and minima increase or decrease with rising of the star? Assume that the sea surface is flat.

pr 14. [IPhO-1990] We wish to study X-ray diffraction by a cubic crystal lattice (see figure *a*). To do this we start with the diffraction of a plane, monochromatic wave that falls perpendicularly on a 2-dimensional grid that consists of $N_1 \times N_2$ slits with separations d_1 and d_2. The diffraction pattern is observed on a screen at a distance L from the grid. The screen is parallel to the grid and L is much larger than d_1 and d_2.



(i) Determine the positions and widths of the principal maximum on the screen. The width is defined as the distance between the minima on either side of the maxima.

(ii) We consider now a cubic crystal, with lattice spacing a and size $N_0a \times N_0a \times N_1a$, where $N_1 \ll N_0$. The crystal is placed in a parallel X-ray beam along the z-axis at an angle θ (see Fig. *b*). The diffraction pattern is again observed on a screen at a great distance $L \gg N_0a$ from the crystal. Calculate the position and width of the maxima as a function of the angle θ for $\theta \ll 1$. What in particular are the consequences of the fact that $N_1 \ll N_0$?

(iii) The diffraction pattern can also be derived by means of Bragg's theory, in which it is assumed that the X-rays are reflected from atomic planes in the lattice. The diffraction pattern then arises from interference of these reflected rays with each other. Show that this so-called Bragg reflection yields the same conditions for the maxima as those that you found in (ii).

(iv) In some measurements the so-called powder method is employed. A beam of X-rays is scatted by a powder of very many, small crystals (Of course the sizes of the crystals are much larger than the lattice spacing, a.) Scattering of X-rays of wavelength $\lambda = 0.15\,\mathrm{nm}$ by Potassium Chloride [KCl] (which has a cubic lattice, see Fig *a*) results in the production of concentric dark circles on a photographic plate. The distance between the crystals and the plate is $L = 0.10\,\mathrm{m}$ and the radius of the smallest circle is $R = 0.053\,\mathrm{m}$. (see Fig *c*). K^+ and Cl^- ions have almost the same size and they may be treated as identical scattering centres.

(v) Calculate the distance between two neighbouring K^+ ions in the crystal.

pr 15. [Est-PhO-2009] A hall of a contemporary art installment has white walls and white ceiling; the walls and the ceiling are lit with a monochromatic green light of wavelength $\lambda = 550\,\mathrm{nm}$. The floor of the hall is made of flat transparent glass plates. The lower surfaces of the glass plates are matte and painted black; the upper surfaces are polished and covered with thin transparent film. A visitor standing somewhere in the room will see circular concentric bright and dark stripes on the floor, centred around himself. A curious visitor investigates the phenomenon and concludes the following: in order to see the largest bright stripes, he needs to lower his viewpoint; the maximal number of observable stripes is $N = 20$. Determine the thickness of the film if the film's coefficient of refraction is known to be $n_0 = 1.4$, and that of the glass plates — $n_1 = 1.6$.

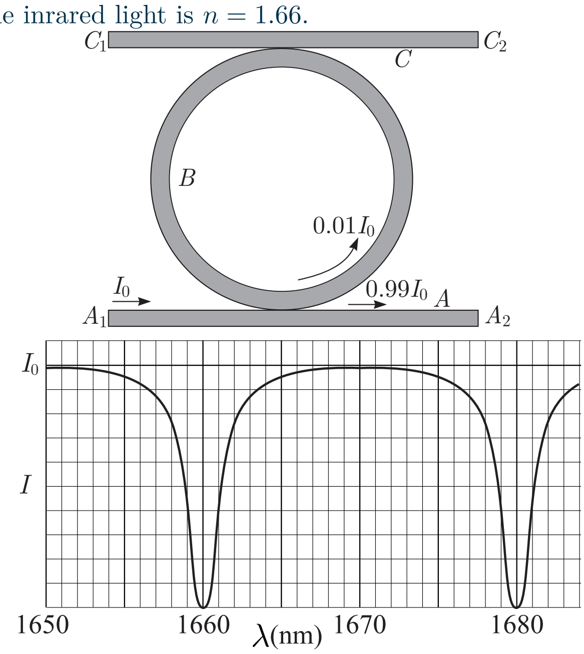
pr 16. [Est-PhO-2002] Circular resonator is a device used in fiber optics; it consists of circular loop made of an optical fibre, coupled to two straight fibres as shown in the figure. Fibre coupling is achieved by bringing the light-conducting cores so close that electromagnetic waves can "tunnel" through the inter-fibre-gap, from one fibre into the other one. In the case of circular resonators, the coupling between the fibres is very weak: if a light pulse propagates along the fibre A from left to right, most of the light energy will pass the coupling point and continue propagation towards A_2, and only a small fraction α of the incident energy "jumps" over to the circular fibre B; let $\alpha = 0.01$. Let us assume the following: (i) all three fibres have identical properties; (ii) these are so called *single-mode fibres*, i.e. light can travel only parallel to the fibre's axis, without "bouncing" between the walls; (iii) the coupling between the fibres B and C is identical to that of between A and C; (iv) monochromatic infrared light of intensity I_0 is being led to the inlet A_1 of the fibre A.

Graph below shows the dependence of the light intensity at the outlet A_2 on its wavelength λ.

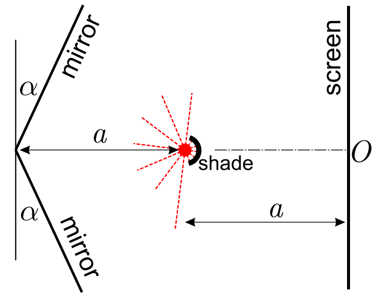
i. Sketch the intensity of light at the outlets C_1 and C_2 as a function of λ.

ii. What is the intensity of light in the fibre B for $\lambda = 1600\,\text{nm}$?

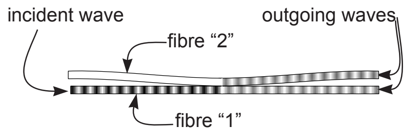
iii. How long is the fibre B? The fibre's coefficient of refraction for the inrared light is $n = 1.66$.



pr 17. [Est-PhO-2004] Screen, two mirrors, and a source of monochromatic light are positioned as shown in figure. Due to a shade, only reflected light from the source can reach the screen. There will be a striped interference pattern on the screen; the distance between the stripes is d. Express the wavelength of the light λ in terms of d and the distance a (see figure). Assume that $a \gg d$.

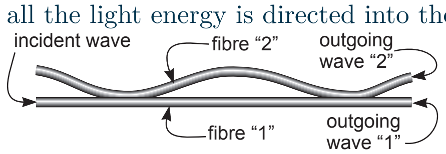


pr 18. [Est-PhO-2001] In fibre optics, devices called equal ratio splitters are often used: these are devices where two optical fibres are brought into such a contact that if an electromagnetic wave is propagating in one fibre, at the contact point it splits into two equal amplitude waves, travelling in each of the fibres, see figure.



1. Show that if an equal ratio splitter splits an electromagnetic wave into two, after the contact point, there is a phase shift of $\frac{\pi}{2}$ between the two waves. *Hint:* use the energy conservation law; depending on your solution, equality $\sin\alpha + \sin\beta = 2\sin\frac{\alpha+\beta}{2}\cos\frac{\alpha+\beta}{2}$ can be useful.

ii. Consider now two sequentially positioned equal ratio splitters, as shown in the figure below (a device called the Mach-Zehnder interferometer). The optical path difference between the inter-splitter segments of the two fibres is $\Delta = 30\,\mu\text{m}$. Assuming that the wavelength of the incoming monochromatic light varies from $\lambda_1 = 610\,\text{nm}$ to $\lambda_2 = 660\,\text{nm}$, for which wavelengths all the light energy is directed into the fibre "2"?



appendix 1: Mathematically, we can derive this from the superposition principle (Maxwell Eqns are linear, hence superposition principle holds: any linear combination of solutions is also a solution), and from a branch of mathematics called the Fourier analysis. The latter states that any function of x can be represented as a sum of sinusoidal functions: $f(x) = \int f_k e^{ikx} dk$. Assume that for $t = 0$, $\vec{E}(z, t = 0) = \vec{e}_x E_\star(z) = \vec{e}_x \int E_k e^{ikz} dk$. Each of the sinusoidal components will evolve in time according to Eq. (3), so that with $\zeta = z - vt$,

$$\vec{E}(z,t) = \vec{e}_x \int E_k e^{ik(z-vt)} dk = \vec{e}_x \int E_k e^{ik\zeta} dk = \vec{e}_x E_\star(\zeta),$$

i.e $\vec{E}(z,t) = \vec{e}_x E_\star(z - vt)$.

Hints

2. Apply method 1.

3. Apply method 1; note that the sum of three vectors of equal length can be zero only if they form an equilateral triangle.

9. Decompose the light after the first polarizer into two components according to the axis defined by the middle polarizer; do the same for the light before the final polarizer.

Answers

2. $\varphi_{\min} = \arcsin\left[\frac{\lambda}{a}(n \pm \frac{1}{3})\right]$; $\varphi_{\max} = \arcsin\left(n\frac{\lambda}{a}\right)$

9. $I = \frac{1}{2} I_0 \cos^2\alpha \sin^2\alpha = \frac{1}{8}\sin^2 2\alpha$.

10. The short-wavelength oscillations on the graph are due to the diffraction on the film, therefore the local maximum condition is $2dn = \lambda N = cN/\nu$. So, $2dn\nu = cN$ and $2dn(\nu+\delta)\nu = c(N+1)$, hence $2dn\delta\nu = c$ and $d = c/2n\delta\nu$. In order to measure the distance between two maxima more precisely, we take a longer frequency interval , e.g. $\Delta\nu = 80\,\mathrm{THz}$ and count the number of maxima between them, $m \approx 34$. Consequently, $\delta\nu = \Delta\nu/m \approx 2.35\,\mathrm{THz}$, and $d \approx 50\,\mu\mathrm{m}$

11. **(i)** First we need to find the angle after the refraction β: For small incidence angles we find approximately $\beta = \alpha/n$. In the liquid, the wavelength is decreased n times: $\lambda' = \lambda/n$. The requested wavelength can be found as the distance between the lines connecting the intersection points of the equal phase lines of the two beams. Alternatively (and in a simpler way), it is found as the difference of the two wavevectors: $k' = k\beta$, where $k = 2\pi/\lambda' = 2\pi n/\lambda$ is the wavevector of the incident beams. So, $\Delta = 2\pi/k' = \lambda/\alpha \approx 7{,}4\,\mu$m.

(ii) The scattered light fluctuates due to the motion of the scattering particles; the frequency is $\nu = v/\Delta = v\alpha/\lambda$. There is no way to determine the direction of the flow, but the modulus is obtained easily: $v = \nu\lambda/\alpha \approx 0.37\,\mathrm{m/s}$.

(iii) The spatial structure of the interference pattern remains essentially unchanged (the wavelength difference is negligible). However, the pattern obtains temporal frequency $\delta\omega = \delta(c/\lambda) \approx c\delta\lambda/\lambda^2$. The velocity of the interference pattern $u = \Delta\delta\omega = \frac{c}{\alpha}\frac{\delta\lambda}{\lambda}$. If the fluid speed is $v \approx 0.37\,\mathrm{m/s}$, then the relative speed of the pattern and the fluid is $\nu' = \frac{c}{\alpha}\frac{\delta\lambda}{\lambda} \pm v$, depending on the direction of the flow (in both cases, $\nu' \approx 740\,\mathrm{kHz}$). So, the output frequency allows us to determine the flow direction as long as we can be sure that the interference pattern velocity is larger than the flow velocity.

12. **(i)** The light flux density decreases inversely proportionally to the square of the distance, therefore $w_1 = w_0 R_p^2/L_p^2$, where R_p is the solar radius, and L_p — the solar distance. Due to $\phi = 2R_p/L_p$, we obtain $w_1 = w_0\phi^2/4$.

(ii) The previous result can be applied to the star flux density, which is $q^{-2}w_1$; hence $P_2 = \frac{1}{4}\pi D^2 w_1 q^{-2} = w_0\pi(\phi D/4q)^2$.

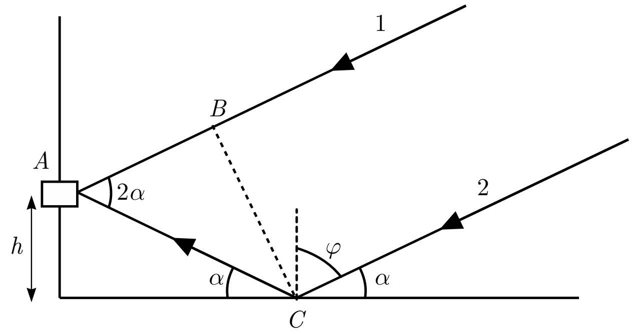
(iii) The paper surface area S radiates towards the lens of the telescope the power $P_3 = w_1\alpha S(\frac{\pi}{4}D^2/L^2)$, where L is the telescope distance. The image of this piece of paper has size $s = SF^2/L^2$; thus, $w_3 = P_3/s = w_1\alpha(\frac{\pi}{4}D^2/F^2) = w_0\alpha\pi(\phi D/4F)^2$.

(iv) The angular distance of the first diffraction minimum (using the single slit approximation — circle is actually not a slit) is λ/D. Hence, the bright circle radius can be estimated as $\delta = F\lambda/D$. Consequently, $w_2 = P_2/\pi\delta^2 = w_0(\phi D^2/4qF\lambda)^2$.

(v) $k = (w_2+w_3)/w_3 = 1+(\alpha\pi)^{-1}(D/\lambda q)^2 \approx 4$ (assuming $\lambda \approx 500\,\mathrm{nm}$).

(vi) $k-1 \sim 1$ (or $k-1 > 1$) means that the star can be easily seen (as is the case for the telescope); $k-1 \ll 1$ means that the star cannot be seen (for the eye, $k-1 \approx 1\cdot 10^{-4}$).

14. **(i)** The signal, registered by the detector A, is result of the interference of two rays: the ray 1, incident directly from the star and the ray 2, reflected from the sea surface (see the figure)



The phase of the second ray is shifted by π due to the reflection by a medium of larger refractive index. Therefore, the phase difference between the two rays is:

$$\Delta = AC + \frac{\lambda}{2} - AB = \frac{h}{\sin\alpha} + \frac{\lambda}{2} - \frac{h}{\sin\alpha}\Big)\cos(2\alpha) =$$
$$= \frac{\lambda}{2} + \frac{h}{\sin\alpha}[1-\cos(2\alpha)] = \frac{\lambda}{2} + 2h\sin\alpha$$

The condition for an interference maximum is:

$$\frac{\lambda}{2} + 2h\sin\alpha_{\max} = k\lambda \quad \text{or}$$
$$\sin\alpha_{\max} = (k-\frac{1}{2})\frac{\lambda}{2h} = (2k-1)\frac{\lambda}{4h}$$

where $k = 1, 2, 3, \ldots, 19$. (the difference of the optical paths cannot exceed $2h$, therefore k cannot exceed 19).

The condition for an interference minimum is:

$$\frac{\lambda}{2} + 2h\sin\alpha_{\max} = (2k+1)\frac{\lambda}{2} \Rightarrow$$
$$\sin\alpha_{\max} = \frac{k\lambda}{2h}$$

where $k = 1, 2, 3, \ldots, 19$.

(ii) Just after the rise of the star the angular height α is zero, therefore the condition for an interference minimum is satisfied. By this reason just after the rise of the star, the signal will increase.

(iii) If the condition for an interference maximum is satisfied, the intensity of the electric field is a sum of the intensities of the direct ray E_i and the reflected ray E_r, respectively: $E_{\max} = E_i + E_r$.

Because

$$E_r = E_i\frac{n-\cos\varphi}{n+\cos\varphi},$$

then

$$E_{\max} = E_i \left(1 + \frac{n - \cos\varphi}{n + \cos\varphi}\right),$$

From the figure it is seen that $\varphi_{\max} = \frac{\pi}{2} - \alpha_{\max}$, we obtain

$$E_{\max} = E_i \left(1 + \frac{n - \sin\alpha_{\max}}{n + \sin\alpha_{\max}}\right) = E_i \frac{2n}{n + \sin(2\alpha_{\max})}. \qquad (4)$$

At the interference minimum, the resulting intensity is:

$$E_{\min} = E_i - E_r = E_i \frac{2\sin\alpha_{\min}}{n + \sin\alpha_{\min}}. \qquad (5)$$

The intensity I of the signal is proportional to the square of the intensity of the electric field E, therefore the ratio of the intensities of the consecutive maxima and minima is:

$$\frac{I_{\max}}{I_{\min}} = \left(\frac{E_{\max}}{E_{\min}}\right)^2 = \frac{n^2}{\sin^2\alpha_{\min}} \frac{(n + \sin\alpha_{\min})^2}{(n + \sin\alpha_{\max})^2}. \qquad (6)$$

Using the eqs. (2) and (3), the eq. (6) can be transformed into the following form:

$$\frac{I_{\max}}{I_{\min}} = \frac{4n^2h^2}{k^2\lambda^2} \left[\frac{n + k\frac{\lambda}{2h}}{n + (2k-1)\frac{\lambda}{4h}}\right]^2.$$

Using this general formula, we can determine the ratio for the first maximum ($k = 1$) and the next minimum:

$$\frac{I_{\max}}{I_{\min}} = \frac{4n^2h^2}{\lambda^2} \left(\frac{n + \frac{\lambda}{2h}}{n + \frac{\lambda}{4h}}\right)^2 = 3.10^4$$

(iv) Using that $n \gg \lambda$, from the Eq. two lines above it follows :

$$\frac{I_{\max}}{I_{\min}} = \frac{4n^2h^2}{k^2\lambda^2}.$$

So, with the rising of the star the ratio of the intensities of the consecutive maxima and minima decreases.

14.

Consider first the x-direction. If waves coming from neighbouring slits (with separation d_1) traverse paths of lengths that differ by:

$$\Delta_1 = n_1.\lambda$$

where n_1 is an integer, then a principal maximum occurs. The position on the screen (in the x-direction) is:

$$x_{n_1} = \frac{n_1.\lambda.L}{d_1}$$

since $d_1 << d_2$.

The path difference between the middle slit and one of the slits at the edge is then:

$$\Delta_{(\frac{N_1}{2})} = \frac{N_1}{2}.n_1.\lambda$$

If on the other hand this path difference is:

$$\Delta_{(\frac{N_1}{2})} = \frac{N_1}{2}.n_1.\lambda + \frac{\lambda}{2}$$

then the first minimum, next to the principal maximum, occurs. The position of this minimum on the screen is given by:

$$x_{n_1} + \Delta x = \frac{\left(\frac{N_1}{2}.n_1.\lambda + \frac{\lambda}{2}\right)L}{\frac{N_1}{2}.d_1} = \frac{n_1.\lambda.L}{d_1} + \frac{\lambda.L}{N_1.d_1}$$

$$\rightarrow \quad \Delta x = \frac{\lambda.L}{N_1.d_1}$$

The width of the principal maximum is accordingly:

$$2.\Delta x = 2.\frac{\lambda.L}{N_1.d_1}$$

A similar treatment can be made for the y-direction, in which there are N_2 slits with separation d_2. The positions and widths of the principal maximal are:

$$(x_{n_1}, y_{n_2}) = \left(\frac{n_1.\lambda.L}{d_1}, \frac{n_2.\lambda.L}{d_2}\right)$$

$$2.\Delta x = 2.\frac{\lambda.L}{N_1.d_1} \; ; \qquad 2.\Delta y = 2.\frac{\lambda.L}{N_2.d_2}$$

An alternative method of solution is to calculate the intensity for the 2-dimensional grid as a function of the angle that the beam makes with the screen.

In the x-direction the beam 'sees' a grid with spacing a, so that in this direction we have:

$$x_{n_1} = \frac{n_1.\lambda.L}{a} \qquad \Delta x = 2.\frac{\lambda.L}{N_0.a}$$

In the y-direction, the beam 'sees' a grid with effective spacing a.cos(Θ). Analogously, we obtain:

$$y_{n_2} = \frac{n_2.\lambda.L}{a.\cos(\theta)} \qquad \Delta y = 2.\frac{\lambda.L}{N_0.a.\cos(\theta)}$$

In the z-direction, the beam 'sees' a grid with effective spacing a.sin(Θ). This gives rise to principal maxima with position and width:

$$y'_{n_3} = \frac{n_3.\lambda.L}{a.\sin(\theta)} \qquad \Delta y' = 2.\frac{\lambda.L}{N_1.a.\sin(\theta)}$$

This pattern is superimposed on the previous one. Since sin(Θ) is very small, only the zeroth-order pattern will be seen, and it is very broad, since $N_1.\sin(\Theta) << N_0$. The diffraction pattern from a plane wave falling on a thin plate of a cubic crystal, at a small angle of incidence to the normal, will be almost identical to that from a two-dimensional grid.

In Bragg reflection, the path difference for constructive interference between neighbouring planes:

$$\Delta = 2.a.\sin(\phi) \approx 2.a.\phi = n.\lambda \quad \rightarrow \quad \frac{x}{L} \approx 2.\phi \approx \frac{n.\lambda}{a} \quad \rightarrow \quad x \approx \frac{n.\lambda.L}{a}$$

Here ϕ is the angle of diffraction.

This is the same condition for a maximum as in section b.

For the distance, √2.a, between neighbouring K ions we have:

$$tg(2\phi) = \frac{x}{L} = \frac{0{,}053}{0{,}1} \approx 0{,}53 \rightarrow a = \frac{\lambda}{2.\sin(\phi)} \approx \frac{0{,}15.10^{-9}}{2\ .\ 0{,}24} \approx 0{,}31\ nm$$

$$K\text{–}K \approx \sqrt{2}.0{,}31 \approx 0{,}44\ nm$$

15. $\frac{N\lambda}{2(n_1 - \sqrt{n_1^2 - 1}} \approx 13\,\mu\mathrm{m}.$