Solution of Physics Cup 2019, Problem No 5

by Johanes Suhardjo

Diffraction grating is usually used to separate polychromatic light (light containing many wavelengths), because the maximum interference for every wavelength depends on the value of the wavelength. The pattern on the screen must be determined by the shape of the diffraction grating and for this problem we have a check-board-like reflective diffraction grating with every cell having size of $20\lambda_0 \times 20\lambda_0$. As it can be easily obtained from the given equation, for every neighbouring cells, the difference of height is $20\lambda_0$ if the two cells are neighbours in the x direction and λ_0 if the two cells are neighbours in the y direction. We can also see that since $|x|, |y| < 5000\lambda_0$, there are in total 500×500 cells on the diffraction grating. Now, a beam of white light which contains all wavelengths from $\lambda_1 = 400$ nm to $\lambda_2 = 700$ nm is coming perpendicularly to the grating. Since the size of every cell is in the same order with the wavelengths of the light, we couldn't use ray optics approximation. Instead, we have to consider the wave nature of the light. This means that although the light beam is coming perpendicular to the grating, the reflected light beam will not only be perpendicular to the grating. There are also beams of light which make a certain angle with the z axis, after being reflected by the grating. One can use Huygen's principle to see why this is the case (the analogy of this is simply one-slit diffraction phenomena. The light beam doesn't only propagate perpendicular to the slit, but there are also beams which make certain angles with the slit). The side surfaces of the grating are assumed to be perfectly absorbing which means that they will give no effect on the pattern except that they will block some of the light (nonetheless, it will be shown that this effect is small).

First, let us consider the effect of the diffraction for every cell. For convenience, let us denote the wave vector $\vec{k} = k_x\hat{x} + k_y\hat{y} + k_z\hat{z}$ of the reflected light beam. Suppose that two light beams are reflected by two points on a cell which are separated by $\Delta\vec{r}$. Just after the reflection, the two beams must have an equal phase, but then one of the beam must travel further, so the two will have phase difference $\phi = \vec{k}.\Delta\vec{r}$ (this is because one of the beam must travel $\Delta\vec{r}.\hat{k}$ further). For 1D diffraction it is easy to obtain using phasors diagram that $E = E_0\frac{\sin(\delta/2)}{\delta/2}$, where E is the resultant electric field, E_0 is the total electric field if they are all adding constructively and δ is the phase difference between the two light reflected by the two ends. Now, for a rectangular cell, we can first sum the electric field for diffraction in the x direction (for a certain value of y), then we sum for all value of y. Therefore, the effect of diffraction by a rectangular cell can be seen as the product of the diffraction effect in each direction, and $E = E_0\left(\frac{\sin\delta_x/2}{\delta_x/2}\right)\left(\frac{\sin\delta_y/2}{\delta_y/2}\right)$. We don't actually need this result, but the important thing is to see that the diffraction can be seen as a product of each x and y direction. For our grating, then we have $\delta_x = k_x 20\lambda_0$ and $\delta_y = k_y 20\lambda_0$. Actually, if k_x or k_y is positive then there will be some light blocked by the side surfaces, so the effective area will no longer be $20\lambda_0 \times 20\lambda_0$, but we will later see that $k_x, k_y << k_z$ and we can conclude that the side surfaces give small effect. Also, as one can observe from the phasors diagram, the diffraction minima occur when δ is an integer multiple of 2π because the phasor diagram will be a complete circle and therefore the resultant electric field is zero.

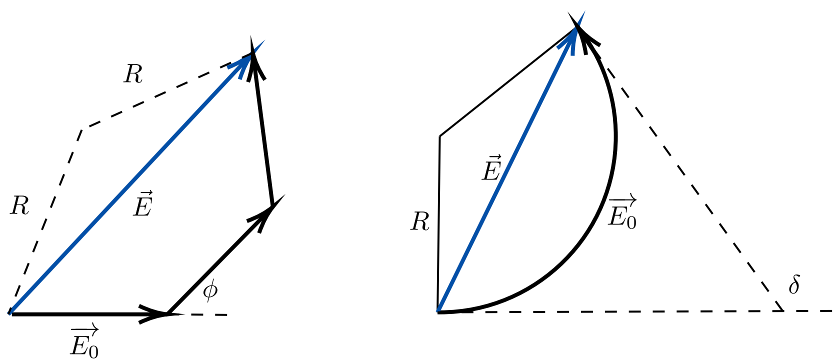

Figure 1: Left figure is phasors diagram for interference of $N = 3$ sources, each having equal magnitude of electric field and phase difference with the neighbour. The right figure shows the phasors diagram for diffraction, the shape is a circle because each small sources can be seen as equally long small phasors and since the change of phase is continuous it makes a circle arc. Note that for interference case $\vec{E}_0$ is the electric field of a source while for the diffraction case it is if all sources add constructively.

After we consider the diffraction effect of one cell, we can now consider the effect of interference between the cells on the grating. Each cell can be represented by a phasor and likewise, we can sum the x direction contribution first, then summing for all y. Hence, the effect of interference can also be seen as a product of each x and y interference. The phase difference between two neighbouring cells can be calculated as follows. Just after being reflected by the two cells, the light beam from the lower cell has traveled Δz more than the other beam. Next, it will also need to travel $\Delta\vec{r}.\hat{k}$ more than the other beam, with $\Delta\vec{r} = \Delta z\hat{z} + \Delta\rho\hat{\rho}$. So, the phase difference is $\phi = k\Delta z + \vec{k}.\Delta\vec{r}$. For two cells neighbouring each other in the x direction, $\Delta z = 20\lambda_0$ and $\Delta\rho = \Delta x = 20\lambda_0$. Thus, $\phi_x = (k + k_z + k_x)20\lambda_0$. Likewise, for two cells neighbouring in the y direction, $\phi_y = (k + k_z)\lambda_0 + k_y 20\lambda_0$. Since there are $N \times N = 500 \times 500$ cells, from the phasor diagram we can see that the interference minima happen when $N\phi = m2\pi$ with m an integer. Notice however that when $\phi = n2\pi$, all the beams add constructively. So, the minima are when m is not an integer multiple of N, otherwise it will be a maximum. Now, we want the brightest spot for every λ. To be clear, after this, when diffraction word is used it refers to the diffraction by each cell alone and interference word will be used when considering the effect of interference between the cells on the grating. The maximum position from the diffraction effect is clearly at point $(0, 0)$ because then the light beam will be reflected perpendicularly. However, this may not be the brightest spot because there is phase difference due to Δz. The brightest spot can't also be too far from $(0, 0)$, because the diffraction effect will greatly reduce the intensity. To determine the position of the brightest spot, let us first write the phase difference for every contribution. For a position (x, y) on the screen we have $k_j = k\frac{j}{s}$ with $j = x, y, z$ and $s = \sqrt{x^2 + y^2 + z^2}$. Now, let us assume that $s \approx z = 50$ cm, we will justify this approximation by showing that our points of interest have positions with $x, y << z$. The first dark spot due to diffraction effect is when $\delta = 2\pi$, so the positions of the first order dark spots are $|x| = x_d = \frac{\lambda}{20\lambda_0}z \approx 2.5$ cm and $|y| = y_d = \frac{\lambda}{20\lambda_0}z \approx 2.5$ cm. We can see that $x_d, y_d << z$. On the other hand, the interference pattern is periodic every time ϕ is increased by 2π, so for the x interference, $\phi_x = (k + k_z + k_x)20\lambda_0 \approx k(2 + x/z)20\lambda_0$ is periodic every time x is increased by $\Delta x_i = \frac{\lambda}{20\lambda_0}z$. Likewise, $\phi_y \approx k\lambda_0(2 + 20y/z)$, so it is periodic every $\Delta y_i = \frac{\lambda}{20\lambda_0}z$. Notice that $\Delta x_i, \Delta y_i$ are also the separation between two neighbouring (global) maxima of the interference effect alone and therefore, there must be at least one interference maximum in the region between the two first order diffraction minima. The brightest spot for every λ must then also lie inside this region. Thus, we can limit our attention inside this region ($|x| < x_d, |y| < y_d$) and since $x_d, y_d << z$ our approximation that $s \approx z$ is justified.

Let us now consider the width of the interference maxima. The dark spot due to the interference effect is when $\phi = m\frac{2\pi}{N}$, consequently the positions of these are $x_{id} = \left(\frac{m_x\lambda}{N20\lambda_0} - 2\right)z$ and $y_{id} = \left(\frac{m_y\lambda}{N20\lambda_0} - \frac{2}{20}\right)z$ with m_x and m_y are integers but not a multiple of N. For both the x and y direction, the width of the interference maxima is simply the distance between $m = N + 1$ and $m = N - 1$, so the width of an interference maximum is $l_{ix} = l_{iy} = \frac{\lambda}{10N\lambda_0}z$. Notice that this width is much smaller than that of the diffraction effect, because $N >> 1$. The width of the interference maximum is only $1/500$ of $2x_d$, so the brightest spot for every λ is very close with the position of interference maximum which is the closest to point $(0, 0)$. Between two interference global maxima, there are several numbers of local maxima but since $N >> 1$ the intensity is small compared the global maxima.

We can now start to determine the position of the brightest spot for every λ. The interference maximum is when $\phi = n2\pi$, so the position of the brightest spot is $x_i = \left(n_x\frac{\lambda}{20\lambda_0} - 2\right)z$ and $y_i = \left(n_y\frac{\lambda}{20\lambda_0} - \frac{2}{20}\right)z$, with integer n_x and n_y. We want n_x and n_y such that x_i and y_i are the closest to $(0, 0)$. Let us first consider the y position. To find the brightest spot, consider the following steps. For the same order (the same value of n_y), as λ is getting bigger, y_i will also be getting bigger. When the brightest spot is at $y = y_d/2$, another bright spot having the same intensity will occur at $y = -y_d/2$. This bright spot is the one from smaller value of n_y, and this is true because the distance between two neighbouring maxima is $\Delta y_i = y_d = \lambda/20\lambda_0$. Once λ is increased such that $y_i > y_d/2$ the brightest spot will then be the one from smaller order, because this is closer to $y = 0$ and thus is brighter due to the diffraction effect of each cell on the grating. The brightest spot is then bounded inside the region $|y| < y_d/2$. Using this fact, one can easily obtain $n_y < \frac{2\lambda_0}{\lambda} + \frac{1}{2}$ and since $\lambda_1 \leq \lambda \leq \lambda_2$, it turns out that only $n_y = 2$ and $n_y = 1$ gives the position of interference maximum inside the region of $|y| < y_d/2$ for all range of λ (for $\lambda = \lambda_1$, $n_y = 3$ is actually also inside the region because when $\lambda = \lambda_1$ the brightest spot is at $y_i = \pm y_d/2$). The same thing can be done to determine the x position of the brightest spot and this brightest spot must also be inside the region $|x| < x_d/2$. Now, for the y position, the order changes when $y_i = y_d/2 = z\lambda/40\lambda_0$, so when $n_y = 2$, the order changes when $\lambda = 4\lambda_0/3$. Thus, we have $y_i = \frac{z}{10}\left(\frac{\lambda}{\lambda_0} - 1\right)$ when $\lambda < 4\lambda_0/3$

and $y_i = \frac{z}{10}\left(\frac{\lambda}{2\lambda_0} - 1\right)$ when $\lambda > 4\lambda_0/3$. Therefore, the position of the brightest spot as a function of lambda is

$$x_i = \left(n_x \frac{\lambda}{20\lambda_0} - 2\right) z$$

$$y_i = \left(n_y \frac{\lambda}{20\lambda_0} - \frac{2}{20}\right) z$$

where n_x, n_y is an integer such that $|x_i| < x_d/2$ and $|y_i| < y_d/2$. For the x position, as Δz is bigger, a change in λ will be more significantly affecting the phase difference, and it can be found that n_x varies more than n_y for the given values of λ (from 50 until 29). The outer boundary for the x position can be expressed in terms of y_i, the y position of the brightest spot,

$$|x_{out}| = \frac{x_d}{2} = \frac{z}{40\lambda_0}\lambda = \left(\frac{y}{2n_y} + \frac{z}{20n_y}\right)$$

This can help us in drawing the diagram. Notice that, in deriving the above equations, we make the assumption that the intensity at $x = x_d/2$ is the same with at $x = -x_d/2$ and the intensity at $y = y_d/2$ is the same with at $y = -y_d/2$, however, the side surfaces are blocking some of the light which goes to $x, y > 0$ on the screen. Since $z \approx s$, the angle is small and therefore the effect is negligible (actually, the effect of the blocking is first order especially for the x direction but it will be cancelled by the diffraction, see appendix). Finally, as it can be easily checked when $\lambda = \lambda_1 = 400$ nm the brightest spot is at $y = -1$ cm (as well as $y = 1$ cm) and $x = 0$ cm with $n_x = 50$ and when $\lambda = \lambda_2 = 700$ nm, the brightest spot is at $x = 1.5$ cm, $y = -1.5$ cm with $n_x = 29$. The brightly illuminated region is sketched below.

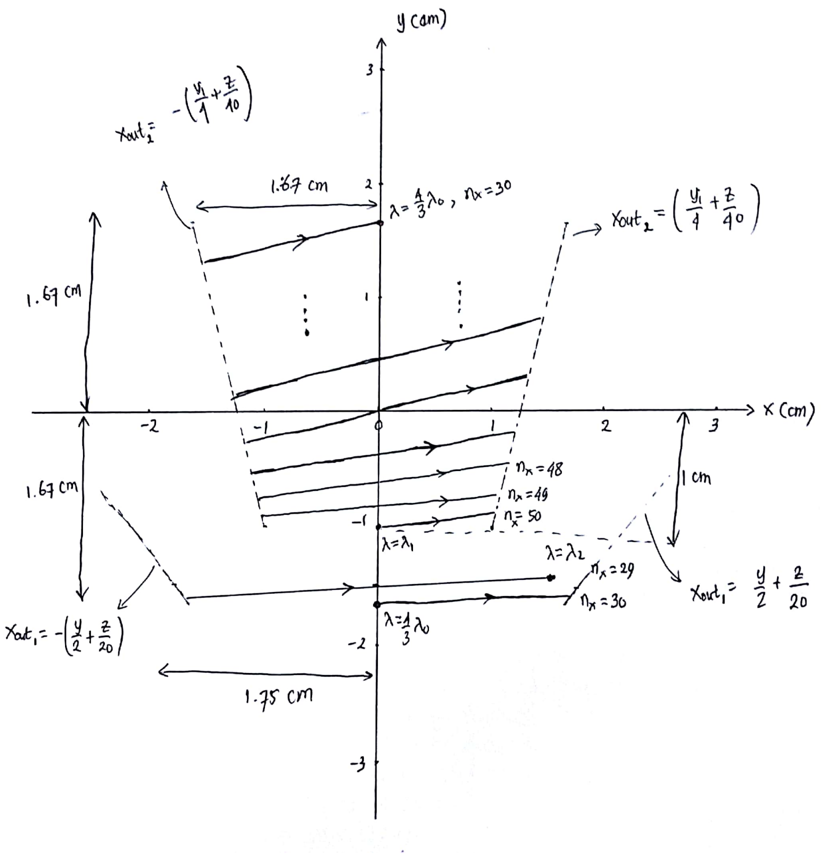

Figure 2: Handwritten sketch showing the brightly illuminated region on the screen. The arrow is showing the direction of increasing wavelength.

Appendix

The intensity as a function of x and y on the screen can be obtained exactly, although this is not so needed as the above approximation is very good. From the phasors diagram, we can obtain that the resultant electric field for the interference effect. One can see from the geometry, that $R = \frac{E_0}{2\sin(\phi/2)}$, so we have $E = 2R\sin(N\phi/2) = E_0 \sin(N\phi/2)/\sin(\phi/2)$. The intensity is then given by

$$I(x,y) = I_0 \left(\frac{\sin(\delta_x/2)}{\delta_x/2}\right)^2 \left(\frac{\sin(\delta_y/2)}{\delta_y/2}\right)^2 \left(\frac{\sin(N\phi_x/2)}{\sin(\phi_x/2)}\right)^2 \left(\frac{\sin(N\phi_y/2)}{\sin(\phi_y/2)}\right)^2 \frac{A^2}{A_0^2}$$

where A is the effective area of each cell and $A_0 = 400\lambda_0^2$ is the area of each cell. Here, I_0 is the intensity if light from a cell adds constructively. The effective area, A is equal with A_0 for points $x, y < 0$ (because no light then will be blocked by the side surfaces). For points with $x, y > 0$, some of the light will be blocked, and if $x > z$ and $y > 20z$ all the reflected light will be blocked. This is very far from our region of interest, because we are only interested in the region $x, y < x_d/2, y_d/2\ z/40$. However, for the x direction, the effective length of each cell is then is around $(1 - 1/40)20\lambda_0$. As we can see, this is still first order. Nevertheless, notice that δ_x also changes. Now, $\delta_x = k_x a$ with a is the effective length, so now the denominator of the term inside the first bracket above cancels the blocking effect. So, the effect of the side surfaces is only inside the δ_x and δ_y which are inside the sin function and it is negligible.