Diffraction and coherent optical signal processing

Physics Laboratory for Master Students in Physics
Sarkadi Tamás

1. Diffraction

1.1. Introduction

Diffraction of light at optical slits and gratings is an interference phenomenon characteristic for waves. The propagation of waves is dictated by so-called *wave equations*, which are differential equations in time and space. The mathematically correct description of wave propagation is rather complex but the *Huygens-Fresnel* principle introduced in the next chapter offers a rather instructive alternative approach to understand on a basic level interference phenomena like refraction and diffraction.

1.2. Basic Theory

Light is an electromagnetic wave consisting of time-varying electric and magnetic fields, which oscillate in time and propagate in space similar to a mechanical wave or sound wave. Unlike mechanical waves electromagnetic waves require no medium and travel, e.g., in vacuum with a constant velocity (c = 299792458 m/s). The fields carry both energy and momentum. In the following it will suffice to consider the electric field $\vec{E}$ alone, which may be considered to play a role similar to the deflection from equilibrium position in a mechanical wave. A *plane wave* with frequency ν and wavelength λ travelling in the positive *x*-direction the electric field is mathematically described by a function

$$E(x,t) = E_0 \cdot \sin\left(2\pi\nu \cdot \left(t - \frac{x}{c}\right)\right) \tag{1.1}$$

This equation tells us that the electric field $E(x,t)$ determined at a certain position x in space oscillates in time with a frequency ν and an amplitude E_0. The argument of the sin-function is often called phase. In three dimensions the points of a fixed phase in the wave of equation (1) form *yz*-planes, a fact which may be understood as motivation for the name. Note that with increasing time the spatial positions of fixed phase propagate into positive *x*-direction with the light velocity c. Because of the periodicity of the sin-function, points separated in *x*-direction by an integer multiple of c/ν have the same field value. Thus the wave field is spatially periodic with period λ=c/ν where λ is the wavelength.

Plane waves, such as the one of equation (1.1), are mathematically very simple but by no means describe all possible wave fields. Much more complicated surfaces of constant phase are possible. In general, electromagnetic waves are generated by a number of oscillating point charges. A single point charge is the source of a *spherical wave*, which spreads with light velocity from the point of the charge into all spatial directions. The wave fronts then form sphere surfaces. The field value of such spherical waves must decrease with increasing distance from the source as indicated in Fig. 1.1.

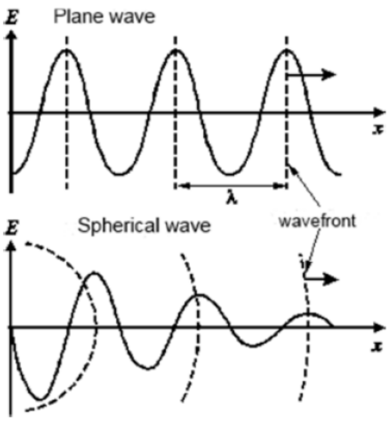


Fig. 1.1: Electric field and wave fronts of a plane wave (top) and a spherical wave (bottom).

If the sources or the wave fronts at some instant are given we may calculated the wave fronts at any later time by the so-called *Huygens' principle*:

- Every point on a wave front serves as the source of a spherical secondary wavelet that spreads out in all directions with a speed equal to the speed of wave propagation.

- The new wave front at a later instant is the surface tangent to all the secondary spherical waves.

Note that only the wave front in forward direction, i.e., in propagation direction, is considered. The wave front simultaneously formed on the back half is disregarded. It was about 150 years later when Fresnel pointed out that Huygens' principle can be extended in order to explain diffraction effects showing that optical apertures do not cast perfect shadows. This extended version of Huygens' Principle is therefore often named *Huygens-Fresnel principle*:

- Every unobstructed point of a wave front at a given instant serves as a source of a spherical secondary wavelet with the same frequency as that of the primary wave. The amplitude of the wave field at any point in forward direction results from the superposition of all these wavelets, considering their amplitudes and relative phases.

It is important to note that superposition here means addition of the electric field values in contrast to addition of intensities. However, for both the human eye and most detectors the light intensity is of relevance, which is proportional to the square of the electric field:

$$I_{\mathrm{Ligt}} \propto |E|^2 \tag{1.2}$$

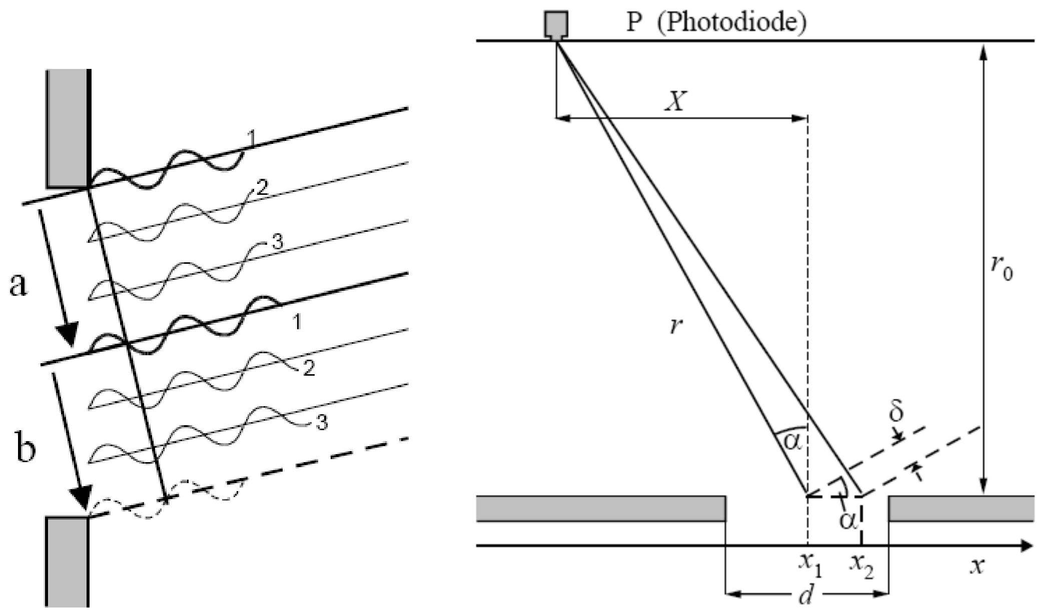
The intensity of the light field at the point of observation arises from the square of the electric field that results from summing up the electric fields of all wavelets. A second crucial point is that the extended version explicitly allows superposition of spherical wavelets with different radii. Considering the phase[1] of the wavelets suffices.

In this laboratory exercise interference phenomena will be studied in the limit of *Fraunhofer* diffraction (as opposed to *Fresnel* diffraction). This means that, firstly, the beam impinging on the diffracting obstacle can be assumed to be a plane wave and, secondly, the distance between the diffracting object and the point of observation is very large. In our experiment the first condition is met in good approximation since we use a laser as light source. The second condition is met if the distance between the obstacle and the point of observation is much larger than the extent of the diffracting object, i.e., the width of the slit or the illuminated grid.

In this experiment the wave properties of light will be demonstrated by observing diffraction and interference effects from a single slit, a double slit and a series of slits. We will learn about the properties of diffraction gratings, which are important components in spectroscopic instruments such as monochromators and spectrographs.

1.3. Diffraction from a single slit

Here we consider a monochromatic plane wave incident on a long narrow slit. According to geometric ray optics the transmitted beam would have the same cross section as the slit, but what we indeed observe is a drastically different diffraction pattern. The main properties of the diffraction pattern can be understood with the aid of Fig. 1.2. For the case of Fraunhofer diffraction, where the distance *r* between the slit and the point of observation P is very large compared to the slit width *d*, all rays origination from the slit can be considered to be parallel in good approximation. To calculate the resultant intensity with Huygens' principle we consider the optical path length difference δ between wavelets originating at two different points x_1 and x_2 within the slit. The difference in path length to the point P is $\delta = \frac{d \sin\alpha}{2}$ where *d* is the slit width and α the angle between the optical axis of the slit (line normal to the aperture) and a line from the slit to the point of observation P.



(I) (II)

Fig. 1.2: I) Interference of wavelets originating from different positions within a single slit. The distance of the point of observation P is so far that the rays can be considered parallel. The numbers indicate wavelets that interfere destructively at P. II) Geometry of an optical slit. The positions x_1 and x_2 of two wavelet source points within the slit are marked. The lines connecting these points with the photodiode denote the radii of wavelets interfering at the detector.

A simple geometrical argument explains compellingly why we observe far behind the slit a number of parallel stripes with different intensities. As indicated in Fig. 1.2. (I) we divide the wave front within the slit into two parts *a* and *b* of the same width. At a certain angle each wavelet of the second half will have a partner wavelet in the first half with precisely a half-wavelength optical path difference. This, however, means that at the point of observation all wavelets from the slit add with their partner wavelets to zero because of their 180° phase difference. As a result at this angle the intensity must be at a minimum. The same will occur at any angle, at which we may divide the slit into an even number of slices with each having half a wavelength path difference to the wavelets of the next neighbouring slice. This, finally, results into the condition for angles at which stripes of with vanishing intensity are observed

$$\sin \alpha_{\min} = m \frac{\lambda}{d} \quad \text{with} \quad m = \pm 1, \pm 2, \cdots \tag{1.3}$$

If, on the other hand, the number of slices with wavelets of 180° phase shift is odd, one slice remains unextinguished and an intensity maximum occurs:

$$\sin \alpha_{\max} = \left(m + \frac{1}{2}\right) \frac{\lambda}{d} \quad \text{with} \quad m = \pm 1, \pm 2, \cdots \tag{1.4}$$

The integer number *m* is often called *order* of the maximum. Mathematically the electric field at the point of observation results from a sum of all wavelets originating from any location of a wave front within the slits. This sum turns into an integral since the sources of the wavelets are arbitrarily close to each other. Finally, the intensity at point P is obtained by taking the square of the sum:

$$I_S = I_0 \left(\frac{\sin\left(\frac{\pi \cdot d}{\lambda} \sin \alpha\right)}{\frac{\pi \cdot d}{\lambda} \sin \alpha} \right)^2 \tag{1.5}$$

Here I_0 is the intensity of the maximum observed at angle $\alpha = 0$. Note that the intensity of equation (1.5) indeed vanishes whenever the angle α meets equation (1.3). Furthermore, according to equation (1.5), we expect an intensity maximum (in quite good approximation for $\frac{d \sin \alpha}{\lambda} > 1$ whenever the argument of the sin-function is an odd multiple of π/2 which is in correspondence to equation (1.4).

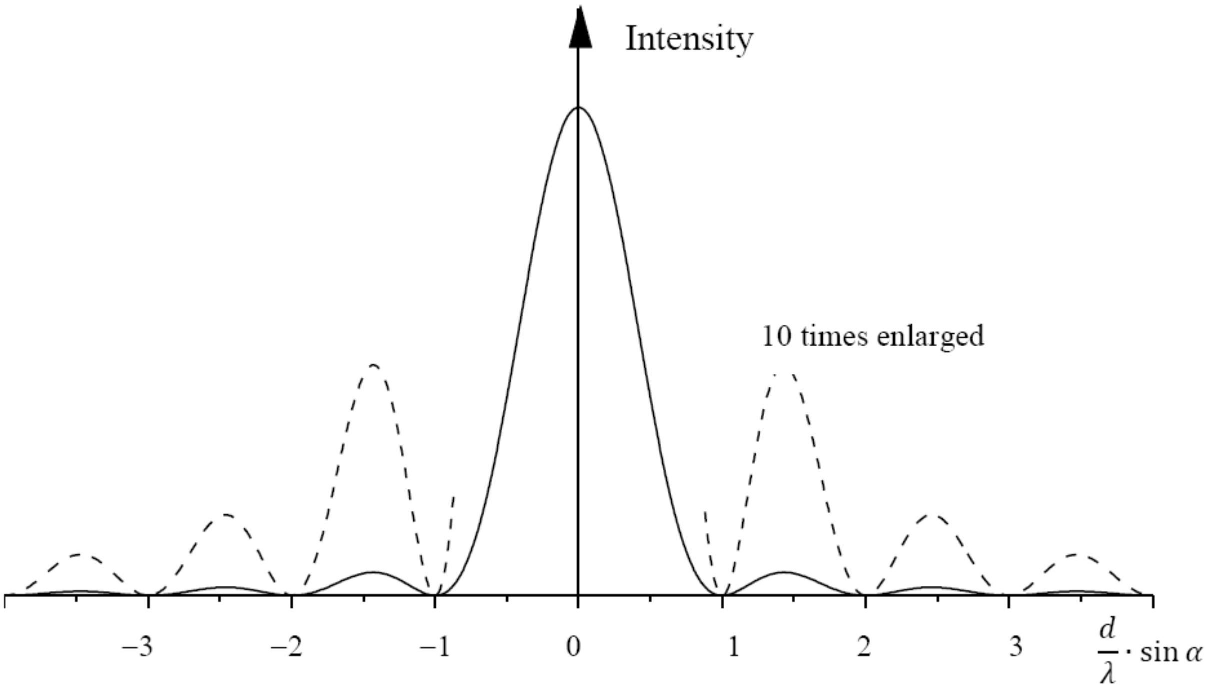


Fig. 1.3. Angular dependence of the intensity pattern resulting from interference at an optical slit. The dashed line depicts the intensity ten times enlarged at angles beyond π.

In Fig. 3 the intensity is depicted as function of the entity $\frac{\pi d \sin\alpha}{\lambda}$ A large main maximum is observed at zero deflection. The intensity of the much weaker side maxima decreases with increasing order *m* according to

$$I_{\max} = \frac{I_0}{\pi^2(|m| + 1/2)^2} \tag{1.6}$$

1.4. The Diffraction Grating

A diffraction grating consists of a large number *N* of identical slits each of width *d* and separated from the next by a distance *D*, as shown in Fig. 1.4.

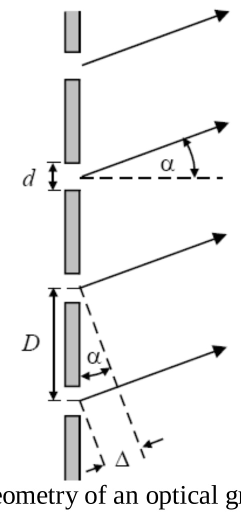


Fig. 1.4: Geometry of an optical grating.

We will first consider the interference of wavelets originating from the slit centers. Of course, the result will then be only a good approximation if the slits widths are extremely small. But we will find an easy way to extend the result to the case of grids with slits of realistic widths. If we assume that the incident light is planar and perpendicular to the grid, the optical path difference of wavelets in neighboring slits is $\Delta = D\sin\alpha$ where the angle α is again taken between the normal to the grating and the line joining the grating and the point of observation P. If this path difference is equal to an integral multiple of wavelengths then all the slits will constructively interfere with each other and a bright spot will be measured at P:

$$\sin\alpha_{\max} = m\frac{\lambda}{D}, \qquad m = 0, \pm 1, \pm 2, \ldots \tag{1.7}$$

These maxima are called principal maxima, because the superposition of the wavelets according to Huygens' principle will turn out to result in additional maxima, which are much weaker though. Mathematically the superposition results in a sum of the fields in the wavelets from all N slits in the grating. To obtain the intensity the result is squared again, which results in:

$$I_G = I_S \left(\frac{\sin\left(\frac{N\cdot\pi\cdot D}{\lambda}\sin\alpha\right)}{\sin\left(\frac{\pi\cdot D}{\lambda}\sin\alpha\right)} \right)^2 \tag{1.8}$$

Extinction occurs at angles α_{min} with

$$\sin\alpha_{\min} = \frac{\ell}{N}\cdot\frac{\lambda}{D}, \qquad \ell = \pm 1, \pm 2, \ldots \text{ und } \quad \ell \neq N \tag{1.9}$$

If the ℓ/N is an integer number, the denominator of I_G in equation (1.8) also vanishes. A mathematical consideration leads to the result that in this case the intensity must be $N^2 I_s$. These maxima occur at angles in accordance with equation (7), and they are much stronger as compared to the intensity observed at any other angle. According to equation (1.9), we expect between principal maxima (N ˜ 1) intensity minima and (N – 2) weak secondary maxima, which strength relative to the principal maxima rapidly decreases with increasing N.
If in each slit we have only one source of a wavelet the factor I_s in equation (1.8) is the intensity of the wavelet. In a real grid we have to consider the finite width of the slit. If different wavelets originating in the same slit are interfering destructively, they will do so in a grating as well. The finite width of the slits is considered in equation (1.8) if the factor I_S contains the intensity pattern of a single slit of the grating. As a result we obtain the interference pattern depicted in Fig. 1.5. In the case depicted in this figure, the width of the slit amounts to 1/5th of the grating period.

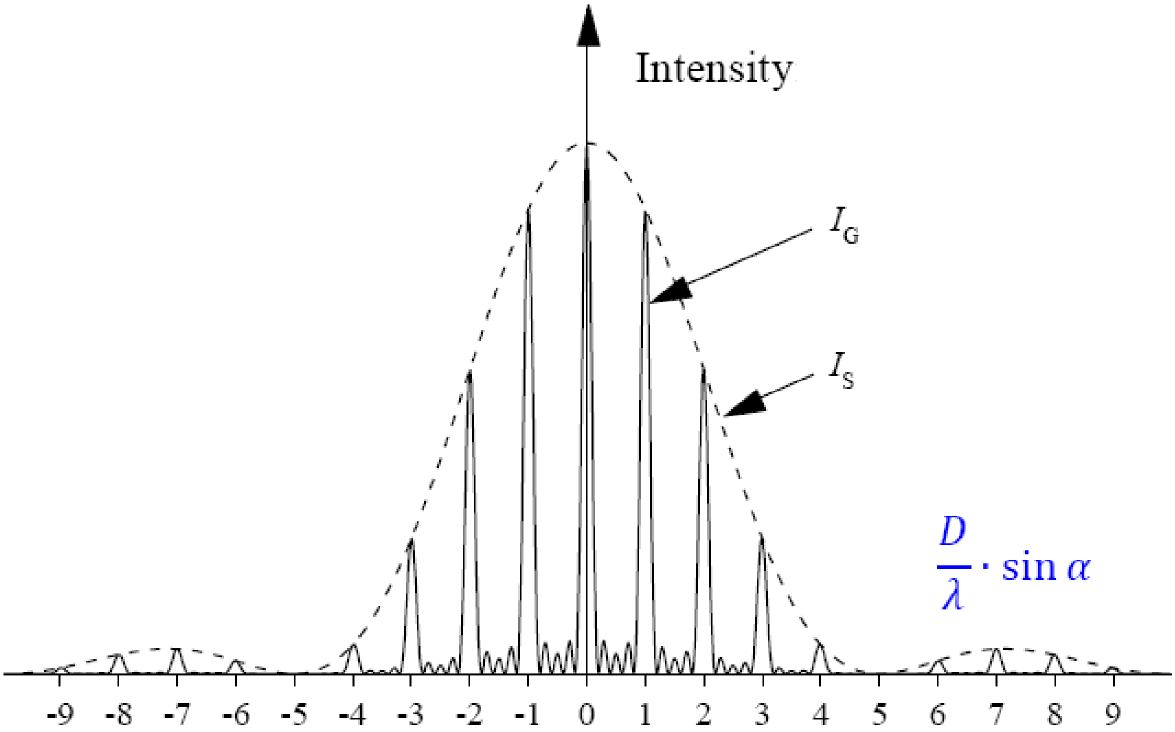


Fig. 1.5: Intensity pattern of an optical grid with N = 5 slits. The ratio between the width of the slits and the period of the grating is $\gamma = 1/5$. The dashed line marked I_S reflects the intensity pattern of a single slit in the grid.

The location of the maxima does not depend on the number of slits. However, they become sharper and more intense as *N* is increased. Because of the wavelength dependence of the angles, at which principal maxima are observed, light consisting of different wavelengths can be spectrally resolved with an optical grating: After diffraction of the light beam at the grating the components with different wavelength are observed at different angles. Note that the re-solving power of the grating, i.e., the resolution at which different wavelengths can be distinguished with the grating, increases with the number of illuminated slits because of the decreasing width of the principal maxima at gratings with increasing *N*.

2. Fourier optics

2.1. Background

Ray optics is a convenient tool to determine imaging characteristics such as the location of the image and the image magnification. A complete description of the imaging system, however, requires the wave properties of light and associated processes like diffraction to be included. It is these processes that determine the resolution of optical devices, the image contrast, and the effect of spatial filters. One possible wave-optical treatment considers the Fourier spectrum (space of spatial frequencies) of the object and the transmission of the spectral components through the optical system. This is referred to as Fourier Optics. A perfect imaging system transmits all spectral frequencies equally. Due to the finite size of apertures (for example the finite diameter of the lens aperture) certain spectral components are attenuated resulting in a distortion of the image. Specially designed filters on the other hand can be inserted to modify the spectrum in a predefined manner to change certain image features.
In this lab you will learn the principles of spatial filtering to (i) improve the quality of laser beams, and to (ii) remove undesired structures from images. To prepare for the lab, you should review the wave-optical treatment of the imaging process (see [1], Chapters 6.3, 6.4, and 7.3).

2.2. Spatial Fourier Transform

Consider a two-dimensional object{ a slide, for instance that has a field transmission functionn *f(x; y)*. This transmission function carries the information of the object. A (mathematically) equivalent description of this object in the Fourier space is based on the object's amplitude spectrum

$$F(u,v) = \frac{1}{(2\pi)^2} \int\int f(x,y) e^{i2\pi ux + i2\pi vy} dx dy, \qquad (2.1)$$

where the Fourier coordinates ($u; v$) have units of inverse length and are called spatial frequencies. Suppose a plane wave of amplitude E_0 impinges on the object. The field distribution immediately behind the object $E(x; y) = f(x; y)E_0$, and the object information is impressed onto the light wave. There are optical processes that can produce the Fourier transform of $E(x; y)$ and the object function, respectively.

2.3. Fourier optical approximation of diffraction

When light (or any other wave) encounters an obstacle, it is diffracted. Diffracted waves are in general characterized by a complicated distribution of amplitude and phase. In certain cases, however, analytical expressions can be obtained for the diffracted field and the diffraction pattern that can be observed on a screen. Consider a rectangular aperture characterized by the transmittance function $f(x; y) = 1$ for $-a/2<x<a/2$ and $-b/2<y<b/2$ where the ($x; y$) plane is the plane of the aperture. It can be shown that the intensity distribution in the far field (Fraunhofer) diffraction pattern is given by

$$I(u,v) = I_0 \frac{\sin^2(2\pi vb/2)}{(2\pi vb/2)^2} \frac{\sin^2(2\pi ua/2)}{(2\pi ua/2)^2} = I_0 \mathrm{sinc}^2(2\pi vb/2)\mathrm{sinc}^2(2\pi ua/2), \quad (2.2)$$

where $u; v$ are the spatial frequencies for the x and y dimensions. For a point $(x_0; y_0)$ in the observing plane at a distance of R_0' from the aperture center, the spatial frequencies are related to the sine of the angle between the normal to the aperture and the line from aperture center to $(x_0; y_0)$ by

$$u = -\frac{x'}{R_0'\lambda}$$
$$v = -\frac{y'}{R_0'\lambda}$$

Note that Eq.(2.2) is simply the absolute value squared of the Fourier transform of $f(x; y)$, i.e $|F(u,v)|^2$ For small diffraction angles, R_0' can be identified with the distance between the diffrracting and observation plane $\sin\alpha \approx \tan\alpha \approx \alpha$. In general, it can also be shown that the far field or Fraunhofer diffraction pattern is the Fourier transform of the field across the diffracting aperture.
Similarly, diffraction from a circular aperture of diameter D, with a transmittance function $f(r,\phi)=1$ for $0<r<D/2$ and $0<\phi<2\pi$, yields

$$I(\rho) = I_0 \left[\frac{2J_1(2\pi\rho(D/2))}{2\pi\rho(D/2)}\right]^2, \quad (2.3)$$

where J_1 is the Bessel function of the first order. The spatial frequency is $\rho = \frac{\sin\theta}{\lambda} \approx \frac{r'}{R_0'\lambda}$ for small diffraction angles θ. The other variables are: D = aperture diameter; $r_0 = r_{disk}$ = radius of Airy disk in question; R_0' = distance from aperture to detection plane; and λ = wavelength. From this intensity pattern, one sees a dependence on the zeroes and extrema of the Bessel function. For example, the distance from the central intensity maxima to the first minimum, r_{disk}, gives the relation

$$1.22 = \frac{\pi D r_{\mathrm{disk}}}{R_0'\lambda} \quad (2.4)$$

while the distance from the central maximum to the next maxima gives

$$1.64 = \frac{\pi D r_{\mathrm{disk}}}{R_0'\lambda} \quad (2.5)$$

Strictly speaking the far field diffraction pattern can only be observed at infinity. By placing a lens after the diffracting aperture the plane at in¯nity is imaged onto the focal plane of the lens. This explains why a lens can perform a Fourier transform.

2.4 Optical spatial filtering

Fourier transform by a lens: Optical spatial fltering is based on the Fourier transform property of a lens (see Fig. 2.1). It is possible to display the two-dimensional spatial frequency spectrum of an object in such a way that individual spatial frequencies can be filtered. This property is illustrated below. The object, in the form of a transparency, is illuminated by a collimated wave. The object is described by a space function *f(x; y)*. In the infinite lens limit, it can be shown [1] that the field distribution in the focal plane of the lens is given by

$$E(x_f, y_f) = \frac{iE_0}{\lambda f} \exp\left\{i\left[\omega t - k(S+f)\right]\right\} \exp\left[i\frac{kr_f^2}{2f^2}(S-f)\right] F(u,v). \tag{2.6}$$

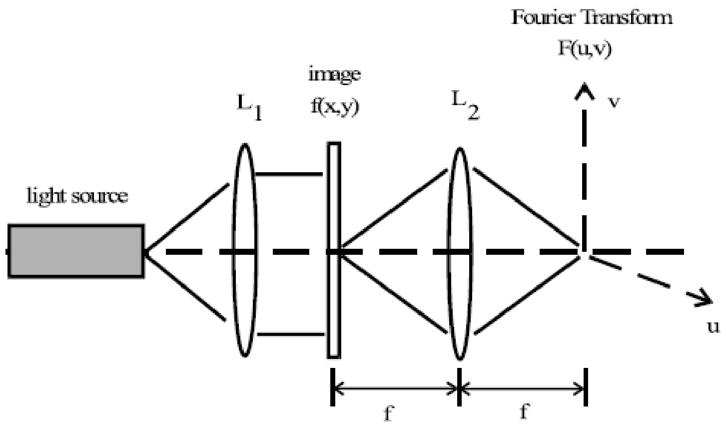


Fig 1: Fourier transform by a lens. L_1 is the collimating lens, L_2 is the Fourier transform lens, *u* and *v* are normalized coordinates in the transform plane.

Here *S* is the object distance, *f* is the focal length of the lens, $r_f^2 = x_f^2 + y_f^2$ are coordinates in the focal plane, *F(u; v)* is the Fourier transform of the object function, $u = -x_f/\lambda f$ and $v = -y_f/\lambda f$. Note, that the observable intensity pattern in the focal plane of the lens, $|E(x_f, y_f)|^2$ does not depend on the object position.

By making a Fourier transform of an image using a lens, it is possible to change the information in amplitude and phase that is supported by this image. For that purpose, a filter is placed in the spatial frequency plane (*u; v*). A second lens, placed after the spatial frequency plane, is used to display the Fourier transform of *F(u; v)T(u; v),* where *T(u; v)* describes the amplitude transmission of the filter. If the filter is removed, the output plane displays an inverted version of the input transparency (object).

The basic optical processor is shown in Fig. 2.2. The object (a transparency) is illuminated by a coherent plane wave. Two identical lenses are used. Ray tracing shows that the system produces an inverted image of the object in the image plane. The first lens produces the Fourier transform of the object in its back focal plane. By manipulating the information in the Fourier transform plane, we obtain a processed (filtered) image."

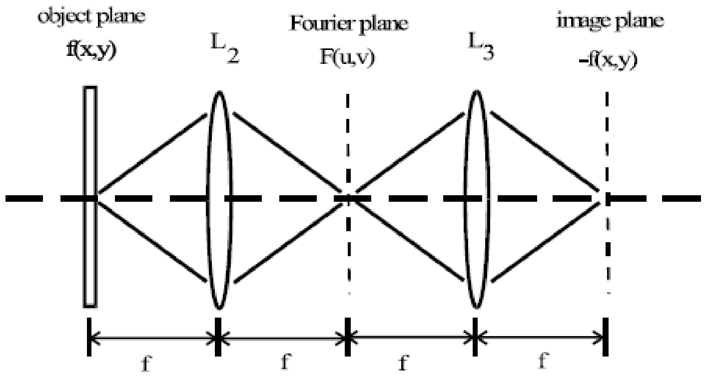


Fig. 2.2. Basic optical filtering system.

Spatial filtering: The filtering consists of altering the image content by placing masks in the Fourier plane. These masks are to affect the amplitude and/or the phase of spectral components.

1. Low pass filtering. A simple kind of filter is the low pass filter frequently used in laser cavities to improve the beam quality. It allows low spatial frequencies to pass through the system while blocking components associated with higher spatial frequencies.

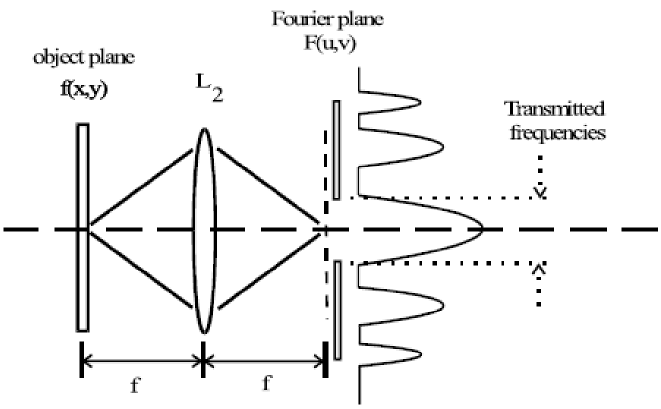


Fig. 2.3: Low pass spatial filter.

2. High pass filter (edge enhancement). Another important filter is the high pass filter (also known as an edge enhancer). A high pass filter consists of a dark spot in the center of the Fourier plane. A high pass filter is used to sharpen photographs. The high pass filter is sketched in Fig. 2.4.

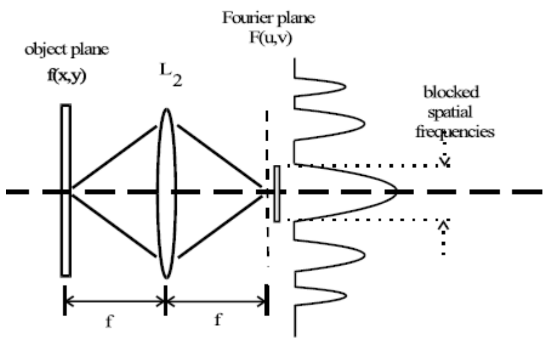


Fig. 2.4: High pass spatial filter

3. Image correction. It is possible to do less trivial filtering. Consider for instance the photograph of a video display. It is possible to eliminate the lines from the picture by inserting two knife edges in the frequency plane. These two edges have to be put at aplace where they cut the +1 and -1 diffraction orders of the grating formed by the lines of the screen. In this way, the picture is passed unchanged and the lines are eliminated.

4. Image correlation. More complicated filters can be made by photography. One interesting application is the comparison of various objects with a reference one, for example, a set of alphanumeric characters. We take one object, say 01, considered as a master. We record it on a photographic plate. Then we take another object and make its FT in the plane where the developed film is carefully placed. If the two objects are identical no light will pass through. If there is a small difference between the two objects, some light will reach the output plane. It is possible this way to test the reproducibility in the manufacturing of complex integrated circuits.

References

[1] Miles V. Klein and Thomas E. Furtak. Optics. John Wiley and Sons, ISBN 0-471-87297-0, New York, 1986.