Optics

Markus Lippitz
September 20, 2023

Contents

Chapter 1
Ray optics

Markus Lippitz
September 18, 2023

Overview

Chapter 2
Gaussian Beams

Markus Lippitz
September 18, 2023

Overview

Chapter 3
Fourier Optics

Markus Lippitz
September 18, 2023

Overview

Fourier transformation simplifies the description of light, especially when it passes through obstacles, as in diffraction. The action of a lens also involves a Fourier transform. This is the field of *Fourier optics*. I will follow chapter 4 of Saleh and Teich, 1991 here. Another good source is Goodman, 2005. Note that books (as Saleh & Teich) from the engineering side of optics use $j = -i = -\sqrt{-1}$ instead of i. Sometimes this j is even written as i, so engineering is the complex conjugate of physics.

We will briefly lay the foundations of Fourier optics and then discuss diffraction and optical Fourier transform through a lens. For our purposes it is sufficient to consider scalar waves, i.e. we ignore the vectorial nature of the electric (or magnetic) field of light and use only a complex scalar value at each point in space to describe light.

Spatial frequencies

Let us start with a plane wave

$$\boldsymbol{U(r) = Ae^{ik\cdot r}} \quad \text{with} \quad \boldsymbol{k = |k| = \frac{2\pi}{\lambda}} \quad . \tag{3.1}$$

We assume that all three components of k are real (*far-field optics* in contrast to *near-field optics*), but the amplitude A might be complex. The wave vector k makes the angles $\Theta_{x,y}$ with the x–z and the y–z plane, respectively, with

$$\sin\Theta_x = \frac{k_x}{k} \quad . \tag{3.2}$$

In the $z = 0$ plane, the field is

$$U(x,y,0) = f(x,y) = A\, e^{2\pi\, i(\nu_x x + \nu_y y)} \tag{3.3}$$

with the *spatial frequencies* ν_x and ν_y

$$\nu_{x,y} = \frac{k_{x,y}}{2\pi} = \frac{1}{\Lambda_{x,y}} \tag{3.4}$$

and the period of the field $\Lambda_{x,y}$ in the x and y direction. And of course all this is related, i.e.,

$$\sin\Theta_x = \frac{k_x}{k} = \lambda\nu_x = \frac{\lambda}{\Lambda_x} \tag{3.5}$$

and similar for the y direction. The assumption of all-real k components makes sure that for all combinations of k_x,k_y, k_z an angle Θ can be found, i.e., the right side of the equation is real and below one in absolute value.

We will almost always make the *paraxial approximation* assuming that the wave vector is roughly parallel to the z-direction, the angles $\Theta_{x,y}$ are thus small, and $k_{x,y} \ll k$. Then we can omit the sine in the last equation and get

$$\Theta_x \approx \frac{k_x}{k} = \lambda\nu_x = \frac{\lambda}{\Lambda_x} \quad . \tag{3.6}$$

What happened here? The combination of all-real k components, i.e., optical far-field, and fixed wavelength λ removes one degree of freedom in the three components of the wave vector. As long as we know the wavelength and we know that the plane wave is nicely propagating, only two real values are enough to fully describe it. These two values could be the angles $\Theta_{x,y}$, or the spatial frequencies $\nu_{x,y}$ or the $\Lambda_{x,y}$.

Transmittance function

A plane wave of amplitude one is traveling in $+z$ direction. At $z = 0$ it is transmitted through a thin optical element with the complex transmittance function $f(x,y)$ with

$$f(x,y) = e^{2\pi\, i(\nu_x x + \nu_y y)} \quad . \tag{3.7}$$

Directly after this plate, the optical field is $U(x,y,0) = f(x,y)$, i.e., the field is modulated by the transmittance function. We know from above that such a field is traveling in the direction given by the $\Theta_{x,y}$ or equally by the spatial frequencies $\nu_{x,y}$. The field is thus diffracted in this direction.[1]

[1] This is not an optical grating yet, as this would change the amplitudes only, i.e., have a real-valued tranmittance function.

In general, if the transmittance function f would have an arbitrary shape, it could be decomposed into a sum of harmonic functions. Each harmonic component would diffract a part of the plane wave into its direction. So when we express f by its Fourier transform F

$$f(x,y) = \mathcal{FT}\left\{F(\nu_x,\nu_y)\right\} = \iint F(\nu_x,\nu_y)\, e^{2\pi\, i(\nu_x x + \nu_y y)}\, d\nu_x d\nu_y \tag{3.8}$$

then we get

$$U(x,y,0) = \iint F(\nu_x,\nu_y)\, e^{2\pi\, i(\nu_x x + \nu_y y)}\, d\nu_x d\nu_y \quad . \tag{3.9}$$

This becomes useful when calculating the field *at any point in space*, i.e., by including the z coordinate:

$$U(x,y,z) = \iint F(\nu_x,\nu_y)\, e^{2\pi\, i(\nu_x x + \nu_y y)}\, e^{ik_z z}\, d\nu_x d\nu_y \quad , \tag{3.10}$$

where k_z now depends on the integrating variables

$$k_z = \sqrt{k^2 - k_x^2 - k_y^2} = 2\pi\left(\frac{1}{\lambda^2} - \nu_x^2 - \nu_y^2\right) \quad . \tag{3.11}$$

Again the requirement of propagating waves entails $\nu_x^2 + \nu_y^2 < 1/\lambda^2$, so not all Fourier components of F play a role.

Transfer function and impulse response

Let us first introduce the concepts with electric circuits such as an RC-filter. One can define a transfer function $H(\omega)$ that relates the frequency spectrum $F(\omega)$ at the input (of the filter) with that at the output

$$G(\omega) = F(\omega) \cdot H(\omega) \quad . \tag{3.12}$$

In time domain, the impulse response $h(t)$ is another description. The signal $f(t)$ at the input results in an output $g(t)$

$$g(t) = \int h(\tau) f(t-\tau) d\tau \quad , \tag{3.13}$$

where causality requires that $h(t)$ is zero for $t < 0$. The interesting point is that not only the signals f and g are connected to their Fourier transforms F and G, but also the transfer function H is the Fourier transform of the impulse response h. A Fourier transform converts a product into a convolution, and vice versa.

Transfer function of free space

We now apply this scheme to spatial frequencies describing a superposition of plane waves. Letting the wave propagate by a distance d from a source plane $f(x,y) = U(x,y,0)$ to a target plane $g(x,y) = U(x,y,d)$, how do the spatial amplitudes F and G relate? Looking at eq. 3.10, we see that it is just the last exponential function that depends on z, but we need to take eq. 3.11 into account. Together we find

$$H(\nu_x, \nu_y) = \exp\left(2\pi\, i\, d \sqrt{\frac{1}{\lambda^2} - \nu_x^2 - \nu_y^2}\right) \quad . \tag{3.14}$$

For spatial frequencies $\nu_x^2 + \nu_y^2 < 1/\lambda^2$, i.e., within a circle of radius $1/\lambda$, the magnitude does not change ($|H| = 1$), only the phase changes. Outside this circle, the magnitude drops exponentially width d, as the square-root becomes imaginary. These waves are called *evanescent waves*, as they do not propagate and only exist in the near-field.

High spatial frequencies ν near $1/\lambda$ are far away from the paraxial approximation. In most cases it is sufficient to restrict ourself to low spatial frequencies $\ll 1/\lambda$. In this case, we can use the *Fresnel approximation* of the transfer function

$$H(\nu_x, \nu_y)_{\text{Fresnel}} = H_0 \exp\left(-2\pi\, i\, d\, (\nu_x^2 + \nu_y^2)\right) \quad \text{with} \quad H_0 = e^{ikd} \quad . \tag{3.15}$$

The term H_0 factors out the trivial phase evolution due to propagation along the optical axis.

When we know the spatial frequency amplitudes F at $z = 0$, then we obtain G at $z = d$ by

$$G(\nu_x, \nu_y) = F(\nu_x, \nu_y) \cdot H(\nu_x, \nu_y) \quad . \tag{3.16}$$

We can Fourier transform the equation to obtain

$$g(x,y) = f(x,y) \otimes h(x,y) \tag{3.17}$$

where $\otimes$ signals a convolution. The impulse response of free space is in the Fresnel approximation

$$h(x,y)_{\text{Fresnel}} \approx h_0 \, \exp\left(ik\frac{x^2+y^2}{2d}\right) \quad \text{with} \quad h_0 = -\frac{i}{\lambda d}\, e^{ikd} \quad . \tag{3.18}$$

Eq. 3.17 means that we get from one plane to the other by convolving each source point with a wave of shape h. This is equivalent to the Huygens principle, where each point should be a source of a spherical wave. When we take the paraxial approximation of a spherical wave we obtain $h(x,y)_{\text{Fresnel}}$.

Optical Fourier transform by propagation

Up to now we used the Fourier transform to simplify description of optical fields. In this section, we will show that the propagation of an optical field by a long enough distance allows to optically 'compute' the Fourier transform. We will find that the field in the target plane $g(x,y)$ is proportional to the Fourier transform F of the field in the source plane.

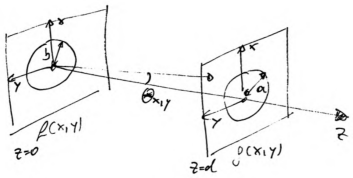

Figure 3.1: Fraunhofer condition

The Fourier components F of the field f in the source plane determine the direction of travel of the plane waves, as we have seen above. The problem is that a plane wave is everywhere in space. We need thus to find a condition for 'far enough' so that the individual pieces of the plane wave have separated enough. We do not only employ the paraxial approximation, i.e., that the wave vectors are not too inclined on the optical axis. The key point is that we also require the size of the source plane to be limited. This leads to the two conditions of the Fraunhofer approximation

$$N_F = \frac{a^2}{\lambda d} \ll 1 \quad \text{and} \quad N_F' = \frac{b^2}{\lambda d} \ll 1 \tag{3.19}$$

where the two N_F are the Fresnel numbers, and a,b are the radius of the relevant and allowed regions in the target and source planes, respectively. d is again the distance between the planes. The Fraunhofer approximation is a more severe restriction than the Fresnel approximation.

We start by writing down the convolution integral of eq. 3.17 in the Fresnel approximation

$$g(x,y) = f(x,y) \otimes h(x,y)_{\text{Fresnel}} \tag{3.20}$$

$$= h_0 \iint f(x',y') \, \exp\left(ik\frac{(x-x')^2+(y-y')^2}{2d}\right) dx'dy' \quad . \tag{3.21}$$

The term $(x-x')^2$ in the exponent of the exponential function is multiplied out into three terms. We keep the mixed terms. Both squared terms can be neglected due to the Fraunhofer approximation. For example we get

$$\exp\left(i\pi\frac{x'^2+y'^2}{\lambda d}\right) \approx 1 \tag{3.22}$$

as $N_F' \ll 1$. The terms without prime vanish due to $N_F \ll 1$. So we have

$$g(x,y) \approx h_0 \iint f(x',y') \, \exp\left(-i2\pi\frac{xx'+yy'}{\lambda d}\right) dx'dy' \quad . \tag{3.23}$$

We now identify the factor $x/\lambda d$ with the spatial frequency ν_x (y similar) and write

$$g(x,y) \approx h_0 F\left(\nu_x, \nu_y\right) = h_0 F\left(\frac{x}{\lambda d}, \frac{y}{\lambda d}\right) \quad . \tag{3.24}$$

When we place a screen g at a distance fulfilling the Fraunhofer condition after a diffracting obstacle f, the interference pattern visible on the screen will be described by the Fourier transform F of f. This simplifies a lot the calculation of single slit, double slit and grating, as typically presented in the introductory optics lecture.

Test yourself

1. Convince yourself that the textbook solution, for example in Demtröder, can be obtained by a Fourier transform.
2. Estimate the required distance so that a typical diffraction grating fulfils the Fraunhofer condition.

Optical Fourier transform by a lens

The distance d required to stay within the Fraunhofer approximation can be prohibitively large. We will see here that a lens is able to shorten the distance between the grating and the screen and still keep the Fourier relation. This explains why spectrometers are not too long, but contain a lens or curved mirror.

From geometrical optics in the paraxial approximation we know already that a lens focuses a beam (angles Θ_x, Θ_y to the optical axis) on a point

$$(x,y) = (f\Theta_x, f\Theta_y) \tag{3.25}$$

in the focal plane, where f describes the focal length of the lens. A lens thus separates plane waves by their propagation direction. As in the beginning of the chapter, we can convert angles into optical frequencies and thus find that the field in the target plane g is proportional the the Fourier amplitude F

$$g(x,y) = \tilde{h}\, F\left(\nu_x, \nu_y\right) = \tilde{h}\, F\left(\frac{x}{\lambda d}, \frac{y}{\lambda d}\right) \quad . \tag{3.26}$$

The remaining question is the prefactor $\tilde{h}$. If it would depend of the spatial coordinates x and y, this would destroy the Fourier transform. To obtain $\tilde{h}$, we multiply the transfer functions of free space for the distance source plane to lens (length d) and lens to target plane (length f). And we need to multiply a transfer function for the lens, as the lens has a thickness profile $t(x,y)$ of a material with a certain index of refraction. All together one obtains[2]

$$\tilde{h}(x,y) = \tilde{h}_0 \exp\left(-i\pi \frac{(x^2-y^2)(d-f)}{\lambda^2 f}\right) \quad \text{with} \quad \tilde{h}_0 = \frac{-i}{\lambda f}\, e^{ik(d+f)} \quad . \tag{3.27}$$

This factor becomes spatially constant when the condition $d = f$ is met. A lens thus performs an optical Fourier transform between is two focal planes. In a spectrometer, the grating sits in the front focal plane of the curved mirror (acting as a lens), the detector in its back focal plane.

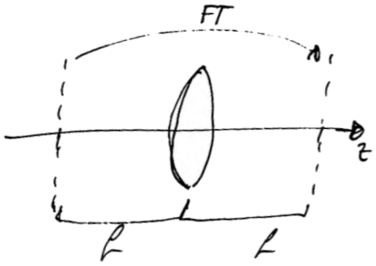

Figure 3.2: Optical Fourier transform by a lens

[2] details in Saleh and Teich, 1991, chapter 4

Spatial filter

In addition to spectrometers, the spatial filter is another important application of a lens as a Fourier transform device. We consider a so-called 4f-system, see Saleh and Teich, 1991. All components are separated by one focal length f: a source plane f, a first lens, a filter plane p, a second lens and a target plane g. Both lenses are identical.

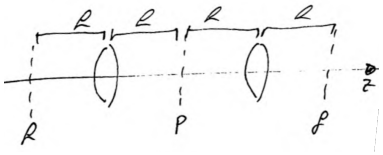

Figure 3.3: A $4f$ system can be used as spatial filter.

Let the transfer function p of the filter plane be $p(x, y) = 1$ for the beginning. Then the first lens Fourier transforms f into F in the filter plane. The filter does nothing and the second lens transforms back F into f, so that we get in the target plane what we started with , i.e., $f = g$. Of course this make the assumption that all plane waves nicely propagate, i.e., the spatial frequencies in f are small enough to cause only propagating plane waves.

The filter plane can be used to modify the Fourier components F. At position x in the p plane, only the Fourier component $\nu_x = x/(\lambda f)$ is present. We can put a mask $p(x, y)$, either just absorbing or with a complex transfer function in the filter plane. The overall transfer function of the 4f-system is then

$$H(\nu_x, \nu_y) = p(\lambda f \nu_x, \lambda f \nu_y) \quad , \tag{3.28}$$

ignoring an overall phase factor for the propagation.

An often used transfer function is a circular aperture. It removes all spatial frequencies above a certain threshold. In this way, one can clean up a laser beam, so that is follows the expected Gaussian profile even after transmission though many non-ideal optical elements.

The inverse filter, i.e. a opaque disc, acts as high-pass filter, increasing the edges in an optical image. A vertical slit lets only pass horizontal features in the image.

References

Goodman, Joseph W. (2005). *Introduction to Fourier optics*. 3. ed. Roberts.

Saleh, Bahaa E. A. and Malvin C. Teich (1991). *Fundamentals of photonics*. New York, NY [u.a.]: Wiley.

Chapter 4
Dielectric Materials

Markus Lippitz
September 18, 2023

Overview

Chapter 5
Birefringence

Markus Lippitz
September 18, 2023

Overview

Chapter 6
Coherence

Markus Lippitz
September 18, 2023

Overview

Chapter 7
Interference

Markus Lippitz
September 18, 2023

Overview

Chapter 8
Quantum Optics

Markus Lippitz
September 18, 2023

Overview

Appendices

Appendix A
Fourier transformation

Markus Lippitz
September 18, 2023

Overview

It is useful and helpful to have an intuitive approach to the Fourier transform. The bottom line is that in experimental physics one rarely needs to actually calculate a Fourier transform. Very often it is sufficient to know a few frequently occurring Fourier pairs and to combine them with simple rules. This is what I want to present here. A very nice and much more detailed presentation can be found in Butz, 2015. I will follow his notation here.

Before we get to Fourier pairs, however, we need to lay down some foundations.

Fourier series: a periodic function and its Fourier coefficients

We first consider everything here in one dimension in time or frequency space with the variables t and $\omega = 2\pi\nu$. Let the function $f(t)$ be periodic in time with period T, i.e.

$$f(t) = f(t+T) \quad . \tag{A.1}$$

Then this can be written as a Fourier series

$$f(t) = \sum_{k=-\infty}^{\infty} C_k \, e^{i\,\omega_k\,t} \quad \text{with} \quad \omega_k = \frac{2\pi\,k}{T} \tag{A.2}$$

and the Fourier coefficients

$$C_k = \frac{1}{T} \int_{-T/2}^{T/2} f(t)\; e^{-i\,\omega_k\,t}\, dt \quad . \tag{A.3}$$

Note the negative sign in the exponential function in contrast to the equation before. For reel-valued functions $f(t)$, 'opposite' C_k are conjugate-complex, so $C_k = C^{\star}_{-k}$. For $k < 0$ the frequencies ω_k are negative, but this is not a problem.[1] Thus, the zeroth coefficient C_0 is just the time average of the function $f(t)$.

[1] One could alternatively require $k \geq 0$ and apply a $\sin$ and $\cos$ series.

An arbitrary function and its Fourier transform

Now we remove the restriction to periodic functions $f(t)$ by letting the period T go to infinity. This turns the sum into an integral and the discrete ω_k become continuous. Thus

$$F(\omega) = \int_{-\infty}^{+\infty} f(t)\, e^{-i\omega\, t}\, dt \tag{A.4}$$

$$f(t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} F(\omega)\, e^{+i\omega\, t}\, d\omega \quad . \tag{A.5}$$

Here, the first equation is the forward transformation (minus sign in the exponent), and the second is the reverse transformation (plus sign in the exponent). The symmetry is broken by the 2π. But this is necessary if one wants to keep $F(\omega = 0)$ as mean[2]. Alternatively, we could formulate all this with ν instead of ω, but then we would have a 2π in many more places, though not before the integral.

[2] $F(0) = \int f(t)\, dt$ without $1/T$ in front of it is meant here by Butz as mean!

Sidenote: Delta Function

The delta function can be written as

$$\delta(x) = \lim_{a\to 0} f_a(x) \quad \text{with} \quad f_a(x) = \begin{cases} a & \text{if } |x| < \frac{1}{2a} \\ 0 & \text{other} \end{cases} \tag{A.6}$$

or as

$$\delta(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{+i\, x\, y}\, dy \quad . \tag{A.7}$$

An important property is that the delta function selects a value, i.e.

$$\int_{-\infty}^{+\infty} \delta(x)\, f(x)\, dx = f(0) \quad . \tag{A.8}$$

Important Fourier pairs

It is very often sufficient to know the following pairs of functions and their Fourier transforms. I write them here, following Butz, as pairs in t and ω (not $\nu = \omega/(2\pi)$). In the same way, one could have written pairs in x and k. The important question is whether a 2π appears in the exponential function of the plane wave or not. So

$$e^{i\omega t} \quad \text{and} \quad e^{ikx} \quad \text{, but} \quad e^{i2\pi\nu t} \quad . \tag{A.9}$$

Further, I follow here the convention made above about the asymmetric distribution of the 2π between forward and reverse transformations. If you distribute them differently, then of course the prefactors change. A good overview of many more Fourier pairs in various '2π' conventions can be found in the English Wikipedia under 'Fourier transform'. In their nomenclature, the Butz convention used here is 'non-unitary, angular frequency'.

constant and delta function $f(t) = a$ becomes $F(\omega) = a\, 2\pi\, \delta(\omega)$ and $f(t) = a\, \delta(t)$ becomes $F(\omega) = a$. This is again the asymmetric 2π.

rectangle and sinc The rectangle function of width b becomes a sinc[3], the sinus cardinalis. So from

$$f(t) = \mathrm{rect}_b(t) = \begin{cases} 1 & \text{for} \quad |t| < b/2 \\ 0 & \text{other} \end{cases} \tag{A.10}$$

we get

$$F(\omega) = b\,\frac{\sin \omega b/2}{\omega b/2} = b\,\mathrm{sinc}(\omega b/2) \quad . \tag{A.11}$$

Gaussian The Gaussian function is preserved under Fourier transform. Its width changes into the reciprocal value. So from a Gauss function of area one

$$f(t) = \frac{1}{\sigma\sqrt{2\pi}}\, e^{-\frac{1}{2}\left(\frac{t}{\sigma}\right)^2} \tag{A.12}$$

we get

$$F(\omega) = e^{-\frac{1}{2}(\sigma\,\omega)^2} \quad . \tag{A.13}$$

(two-sided) exponential decay and Lorentz curve From a curve decaying exponentially at both positive and negative times

$$f(t) = e^{-|t|/\tau} \tag{A.14}$$

we obtain the Lorentz curve

$$F(\omega) = \frac{2\tau}{1 + \omega^2\,\tau^2} \quad . \tag{A.15}$$

one-sided exponential decay As a side note, here the one-sided exponential decay

$$f(t) = \begin{cases} e^{-\lambda t} & \text{for} \quad t > 0 \\ 0 & \text{other} \end{cases} \quad . \tag{A.16}$$

It will become

$$F(\omega) = \frac{1}{\lambda + i\,\omega} \tag{A.17}$$

and it is therefore complex-valued. Its magnitude squared is again a Lorentz function

$$|F(\omega)|^2 = \frac{1}{\lambda^2 + \omega^2} \tag{A.18}$$

and the phase is $\phi = -\omega/\lambda$.

One-dimensional point lattice An equidistant chain of points or delta functions remains an equidistant chain under Fourier transform. The distances take the reciprocal value. So from

$$f(t) = \sum_n \delta(t - \delta t\, n) \tag{A.19}$$

we get

$$F(\omega) = \frac{2\pi}{\delta t} \sum_n \delta\left(\omega - n\frac{2\pi}{\Delta t}\right). \quad . \tag{A.20}$$

Three-dimensional cubic lattice A three-dimensional primitive cubic lattice of side length a makes the transitions to a primitive cubic lattice of side length $2\pi/a$. A face-centered cubic lattice with lattice constant a of conventional unit cell is converted to a space-centered cubic lattice with lattice constant $4\pi/a$ and vice versa.

[3] sometimes $\mathrm{sinc}(x) = \sin(\pi x)/(\pi x)$ is defined, especially when ν and not ω is used as conjugate variable.

Theorems and properties of the Fourier transform

In addition to the Fourier pairs, we need a few properties of the Fourier transform. In the following, let $f(t)$ and $F(\omega)$ be Fourier conjugates and likewise g and G.

linearity The Fourier transform is linear

$$a\,f(t) + b\,g(t) \quad \leftrightarrow \quad a\,F(\omega) + b\,G(\omega) \quad . \tag{A.21}$$

shift A shift in time implies a modulation in frequency and vice versa.

$$f(t-a) \quad \leftrightarrow \quad F(\omega)\,e^{-i\omega a} \tag{A.22}$$

$$f(t)\;e^{-i\omega_0 t} \quad \leftrightarrow \quad F(\omega+\omega_0) \quad . \tag{A.23}$$

scaling

$$f(a\,t) \quad \leftrightarrow \quad \frac{1}{|a|}\,F\left(\frac{\omega}{a}\right) \quad . \tag{A.24}$$

convolution and multiplication Convolution is converted into a product, and vice versa

$$f(t) \otimes g(t) = \int f(\zeta)g(t-\zeta)d\zeta \quad \leftrightarrow \quad F(\omega)\,G(\omega) \tag{A.25}$$

and

$$f(t)\,g(t) \quad \leftrightarrow \quad \frac{1}{2\pi}\,F(\omega) \otimes G(\omega) \quad . \tag{A.26}$$

Parseval's Theorem The total power is the same in both time and frequency domain

$$\int |f(t)|^2\,dt = \frac{1}{2\pi}\int |F(\omega)|^2\,d\omega \tag{A.27}$$

time derivatives

$$\frac{d\,f(t)}{dt} \quad \leftrightarrow \quad i\omega\,F(\omega) \quad . \tag{A.28}$$

Example: Diffraction at a double slit

As an example, we consider the Fourier transform of a double slit, which describes its diffraction pattern. The slits have a width b and a center distance d. Thus the slit is described by a convolution of the rectangular function with two delta functions at the distance d

$$f(x) = \mathrm{rect}_b(x) \otimes (\delta(x-d/2) + \delta(x+d/2)) \quad . \tag{A.29}$$

The Fourier transform of the rectangular function is the sinc, that of the delta functions a constant. However, the shift in position causes a modulation in k-space. Thus, the sum of the two delta functions becomes

$$\mathcal{FT}\left\{\delta(x-d/2) + \delta(x+d/2)\right\} = e^{-ikd/2} + e^{+ikd/2} = 2\cos(kd/2) \quad . \tag{A.30}$$

The convolution with the rectangular function passes into a multiplication with the sinc. Together we get

$$\mathcal{FT}\left\{f(x)\right\} = b\frac{\sin(kb/2)}{kb/2}\,2\cos(kd/2) = \frac{4}{k}\,\sin(kb/2)\,\cos(kd/2) \quad . \tag{A.31}$$

The intensity in direction k is then the squared magnitude of this.

Test yourself

1. *Temporal shift* Sketch the amplitude and phase of the FT of a temporal square pulse pulse centred on time zero! What changes if the pulse is shifted to positive times?

2. *Pulse sequence* You wonder what the Fourier transform (magnitude squared) of an infinite sequence of square pulses looks like and start searching for it on the internet. Your fellow student replies that you can "see" it immediately. Sketch the Fourier transform! Explain why you could derive it directly or why you should "see" it!

3. *Light pulse* Think of a "light pulse" as a mathematical construction of an infinitely long cosine oscillation corresponding to the frequency of light. The "pulse" is obtained by multiplying the wave by a time-limited Gaussian pulse envelope (e.g. half-width of 10 light oscillations). Sketch the construction of the Fourier transform in the spectral domain.

Two-dimensional Fourier transformation

We can extend the definition of the Fourier transform to two and more dimensions. The conjugated variables are (x, y) and (k_x, k_y) instead of t and ω. The wave vector $k_i = 2\pi/\lambda_i$ contains the factor 2π as in the angular frequency ω. We define

$$F(k_x, k_y) = \iint_{-\infty}^{+\infty} f(x,y)\, e^{-i(k_x\, x + k_y\, y)}\, dx\, dy \tag{A.32}$$

$$f(x,y) = \frac{1}{(2\pi)^2} \iint_{-\infty}^{+\infty} F(k_x, k_y)\; e^{+i(k_x\, x + k_y\, y)}\, dk_x\, dk_y \quad . \tag{A.33}$$

When we can separate the function $f(x, y)$ into a product of one-dimensional functions, then the Fourier transform is simply the product of the individual Fourier transforms

$$f(x,y) = g(x)\cdot h(y) \quad \leftrightarrow \quad F(k_x, k_y) = G(k_x)\cdot H(k_y) \quad . \tag{A.34}$$

A rectangle of size $a \times b$ is transformed into a product of sinc functions

$$(x,y) = \mathsf{rect}_a(x)\cdot \mathsf{rect}_b(y) \tag{A.35}$$

$$\leftrightarrow \quad F(k_x, k_y) = ab\,\mathsf{sinc}(k_x a/2)\,\mathsf{sinc}(k_y b/2) \quad . \tag{A.36}$$

A special case of this is the rotational symmetric two-dimensional Gaussian function

$$f(x,y) = \frac{1}{2\pi\sigma^2}\, e^{-\frac{x^2+y^2}{2\sigma^2}} \quad \leftrightarrow \quad F(k_x, k_y) = e^{-\frac{\sigma^2}{2}\left(k_x^2 + k_y^2\right)} \quad . \tag{A.37}$$

One important function can not be separated into a product of one-dimensional functions: a disc of radius a

$$f(x,y) = \begin{cases} 1 & \text{for} \quad x^2 + y^2 < a \\ 0 & \text{other} \end{cases} \tag{A.38}$$

is transformed into

$$F(k_x, k_y) = a\,\frac{J_1(\pi\, a\, \rho)}{\rho} \quad \text{width} \quad \rho = \sqrt{k_x^2 + k_y^2} \tag{A.39}$$

and the (cylindrical) Bessel function of the first kind $J_1(x)$

$$J_1(x) = \frac{1}{\pi}\int_0^\pi \cos(\tau - x\sin\tau)\, d\tau \quad , \tag{A.40}$$

which is the cylindrical analogue of a sinc function.

References

Butz, Tilman (2015). *Fourier Transformation for Pedestrians*. 2. ed. Springer. .

Appendix B
Numerical Fourier Transformation

Markus Lippitz
September 18, 2023

Discrete FT: a periodic sequence of values

In particular, if one collects and evaluates measurement data with a computer, then one does not know the measured function $f(t)$ on a continuous axis t, but only at discrete times $t_k = k\,\delta t$, nor does one know the function from $t = -\infty$ to $t = +\infty$. So we have only a finite sequence of numbers f_k as a starting point. Because we do not know the sequence of numbers outside the measured interval we make the assumption that it is periodic. With N measured values the period is $T = N\Delta t$. For simplicity, we also define $f_k = f_{k+N}$ and thus $f_{-k} = f_{N-k}$ with $k = 0, 1, \ldots, N-1$. Thus the Fourier transform becomes[1]

$$F_j = \frac{1}{N} \sum_{k=0}^{N-1} f_k \, e^{-k\,j\,2\pi i/N} \tag{B.1}$$

[1] see Butz, 2015 chap. 4, Horowitz and Hill, 2015, chap. 1.08, 7.20, 15.18

and its inverse transform

$$f_k = \sum_{j=0}^{N-1} F_j \, e^{+k\,j\,2\pi i/N} \quad . \tag{B.2}$$

The definition is again such that F_0 corresponds to the mean. Because of $f_{-k} = f_{N-k}$, the positive frequencies are in the first half of F_j as the frequency increases. After that come the negative frequencies, starting at the 'most negative' frequency and increasing to the last frequency before zero. So the maximum frequency that can be represented is the Nyquist (angular) frequency

$$\Omega_{\text{Nyquist}} = \frac{\pi}{\delta t} \quad . \tag{B.3}$$

This frequency is such that we take two samples per period of the oscillation. Faster oscillations or fewer samples per period cannot be represented. Even with f_{Nyquist} the imaginary part is always zero, because we always sample the sine at the zero crossing.

FFTW

The most used package for numerical Fourier transform is probably FFTW[2].

[2] https://www.fftw.org/

You have to pay attention to the details of the definition. In particular, the prefactors may differ between different packages. In FFTW, the prefactor $1/N$ changes from the forward to the backward transformation, i.e.

$$F_j = \sum_{k=0}^{N-1} f_k \, e^{-k\,j\,2\pi i/N} \tag{B.4}$$

and the inverse Fourier transform

$$f_k = \frac{1}{N} \sum_{j=0}^{N-1} F_j \, e^{+k\,j\,2\pi i/N} \quad . \tag{B.5}$$

In equations, I (and Butz) use mathematical indices (starting from zero). Some programming languages count from one (e.g., Julia).

One helpful thing of FFTW is that is supplies also a frequency axis. As mentioned above, first come the positive frequencies, starting from zero to the maximum, then the most negative frequency, again rising until just before zero. Depending wether the number of samples N is even or odd, it is a little bit of a hassle to calculate the respective frequencies, but FFTW does this for us:

```
fftfreq(5) # gives [0.0, 0.2, 0.4, -0.4, -0.2]
fftfreq(6) # gives [0.0, 0.166, 0.333, -0.5, -0.333, -0.1.66]
```

Test yourself

1. Try yourself the FFT in a language of your choice. The FFT of, say, $[1111]$ should give something like $[4000]$.
2. The inverse FFT is IFFT. Check that it inverts and test how the pre-factors are distributed.

Wrapping & fftshift

Now lets look at the Fourier transform of a cosine. We evaluate the cosine at 8 points:

$$x_n = n\frac{2\pi}{8} \quad \text{with} \quad n = 0 \ldots 7 \tag{B.6}$$

$$f_n = \cos x_n \tag{B.7}$$

$$F = \mathcal{FT}(f) \quad . \tag{B.8}$$

We find that only F_1 and F_7 are different from zero and have the same, real value. Two values must be different from zero because

$$\cos(x) = \frac{1}{2}\left(e^{ix} + e^{-ix}\right) \quad . \tag{B.9}$$

In general, for real values f_n we have

$$F_{N-j} = F_j^\star \quad . \tag{B.10}$$

The position of these two non-zero values is a consequence of the definition of F_k: first come all positive frequencies and then all negative. For a nicer representation it is often better if the frequency zero is not the first element but in the middle between the positive and negative frequencies. frequencies. This we get by `fftshift` or backwards by `ifftshift`.

Test yourself

3. Convince yourself that you understand why it is element 1 and 7 that differs from zero in the example above.

4. Replace the cosine with a sine in this example and explain the result.

Sampling theorem

We need at least two samples per period to describe a function by its Fourier coefficients. The frequencies must be below the Nyquist frequency f_{Nyquist}

$$f_{\text{Nyquist}} = \frac{1}{2\Delta t} \quad . \tag{B.11}$$

The *sampling theorem* states that this is then also sufficient, i.e., we do not lose any detail by sampling. Let $f(t)$ be a bandwidth-limited function, i.e. $F(\omega)$ is different from zero only in the interval $|\omega| \leq \Omega_{\text{Nyquist}}$. Then the sampling theorem[3] applies and gives

[3] for a proof see Butz, 2015, chap. 4.4

$$f(t) \stackrel{!}{=} \sum_{k=-\infty}^{\infty} f(k\Delta t)\, \text{sinc}\left(\Omega_{\text{Nyquist}} \cdot [t - k\Delta t]\right) \quad . \tag{B.12}$$

So it is enough to sample f all Δt. At the times in between, f is completely described by the (infinitely long) sum of the neighbouring values times the sinc.

In measurement technology, therefore, all we need to do is ensure, for example by means of an electrical filter, that all the frequencies of a signal are below Ω_{Nyquist}, and then our digital acquisition of the signal will be is identical to the signal itself. However, if we sample too infrequently, or if there are higher frequencies present, then these too high frequency components will be reflected at the Nyquist frequency and end up at seemingly lower frequencies. This 'aliasing' distorts the signal.

Zero padding

We began with a repeating pattern of numerical values and their Fourier transform. We always picked the length of the sequence in the examples to match an integer multiple of the period. But of course, this isn't feasible in reality. We lack accurate knowledge of the signal's duration. Or sometimes, multiple signals with varying frequencies are important.

The problem is then a truncation error, which leads to artefacts in the Fourier transform. Fig. B.1 shows an example. 12 data points of a cosine with period 8 are sampled. The FFT assumes periodic continuation (thick) which is not the 'true' signal (thin). In this case, the FFT of the data is far from a peak at the original frequencies. The real part is even spectrally constant (see below Fig. B.2)

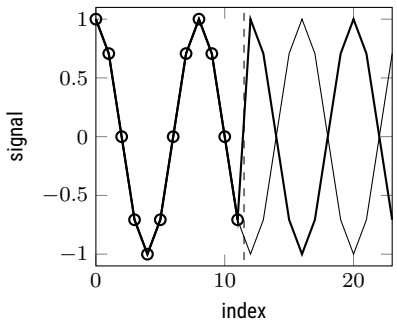

Figure B.1: Clipping a cosine after 1.5 periods

The way out is *zero-padding*. Let our actual measured signal sequence $f(t)$, which we know in the interval $[-T, T]$. Now we pretend that we measured instead

$$g(t) = f(t) \cdot w(t) \tag{B.13}$$

with the window function $w(t)$

$$w(t) = 1 \quad \text{for} \quad -T < t < T \quad \text{other} = 0 \quad . \tag{B.14}$$

Thus we can 'measure' $g(t)$ over arbitrarily long times, because it is is quasi always zero. But the Fourier transform is

$$G(\omega) = F(\omega) \otimes W(\omega) \tag{B.15}$$

with

$$W(\omega) = 2T \, \frac{\sin \omega T}{\omega T} = 2T \, \text{sinc}(\omega T). \quad . \tag{B.16}$$

So we extend our data set on both sides with zeros. The effect is that we convolve the actual Fourier transform of our data set with a sinc whose characteristic width is determined by the actual measurement duration. The frequency resolution does not increase. Rather, a kind of interpolation in Fourier space occurs, which just eliminates the artefacts of the truncation error.

We consider the same data set as above, only we 'extend' it to 10 times the length. This means that the clipping error has less influence and the peak is always at 1 Hz in frequency space. But this does not give more resolution, of course. Peaks that are close to each other cannot be separated by zero-padding, only the position of a peak can be be determined better .

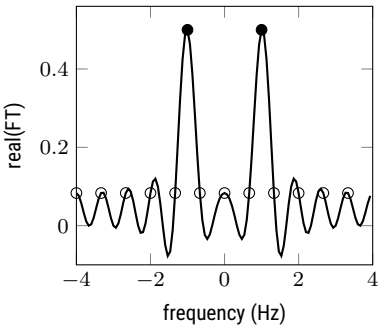

Figure B.2: Zeropadding (line) approaches better the real spectrum (filled symbols) compared to the clipped FT (open symbols).

Windowing

The oscillations in the spectrum in the last example are still artefacts. Actually, one would expect two delta functions at ± 1Hz. They are a consequence of the rectangular window $w(t)$, which leads to the sinc in frequency space. The square-wave window is natural in the sense that we always start and stop measuring. Other window functions[4], however, may be better.They differ the width of the peak and the steepness of the slopes. Unfortunately one must trade one against the other. Interesting parameters are the width of the central peak in frequency space, measured as a -3dB bandwidth, as well as the sideband suppression in [5] dB or its drop in dB/octave.

[4] https://en.wikipedia.org/wiki/Window_function

[5] dB = decibel = $10 \log_1 0x$

Typical window functions are (with $|x| = |t/T| < 1/2$)

$$\text{cosine} = \cos \pi x \tag{B.17}$$

$$\text{triangle} = 1 - 2|x| \tag{B.18}$$

$$\text{Hanning} = \cos^2 \pi x \tag{B.19}$$

$$\text{Hamming} = a + (1-a) \cos^2 \pi x \tag{B.20}$$

$$\text{Gauss} = \exp\left(-\frac{1}{2}\frac{x^2}{\sigma^2}\right) \tag{B.21}$$

$$\text{Kaiser-Bessel} = \frac{I_0(\pi\alpha\sqrt{1-4x^2})}{I_0(\pi\alpha)} \tag{B.22}$$

with the modified Bessel function I_0.

With a window, the measured values are reduced, but the Fourier transform is smoother, because the transition to the zero padding becomes smoother. This makes it possible to recognize in the example the peaks at ± 1Hz even with very few sampled points.

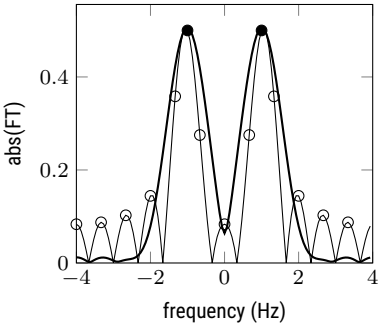

Figure B.3: Zeropadding after windowing (thick) removes the fringes of the un-windowed data (thin) and approaches the true spectrum (solid symbols).

We consider as example[6] a sum of 6 cosine functions with partly very different amplitudes A_l and frequencies f_l:

$$f(t) = \cos\omega t + 10^{-2}\cos 1.15\omega t + 10^{-3}\cos 1.25\omega t + 10^{-3}\cos 2\omega t + 10^{-4}\cos 2.75\omega t + 10^{-5}\cos 3\omega t \tag{B.23}$$

We sample 256 data points at intervals of $\Delta t = 1/8$, i.e. only $8/3 \approx 3$ data points per oscillation of the highest occurring frequency, which is 5 orders of magnitude weaker than the lowest frequency. Nevertheless, this peak can be found with a suitable window and zero-padding.

[6] from Butz, 2015, chapter 3.10

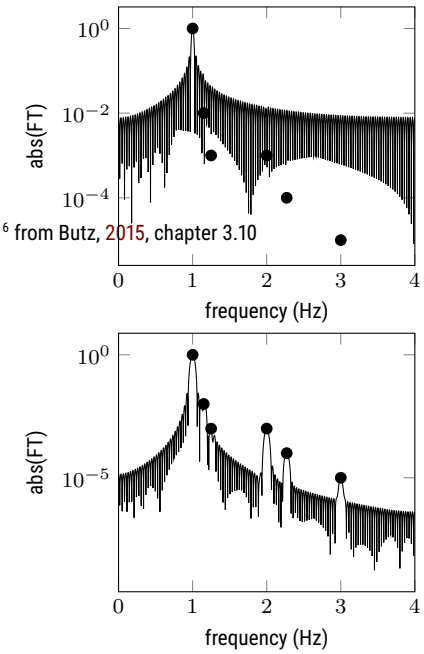

Figure B.4: Without windowing (top), only the main signal component is recovered. A Hanning window (bottom) allows to find even signals 10^{-5} below the main component.

References

Butz, Tilman (2015). *Fourier Transformation for Pedestrians*. 2. ed. Springer. .

Horowitz, Paul and Winfield Hill (2015). *The art of electronics*. Third edition. New York, NY: Cambridge University Press.

Bibliography

Butz, Tilman (2015). *Fourier Transformation for Pedestrians*. 2. ed. Springer. .

Goodman, Joseph W. (2005). *Introduction to Fourier optics*. 3. ed. Roberts.

Horowitz, Paul and Winfield Hill (2015). *The art of electronics*. Third edition. New York, NY: Cambridge University Press.

Saleh, Bahaa E. A. and Malvin C. Teich (1991). *Fundamentals of photonics*. New York, NY [u.a.]: Wiley. .