1
Preliminaries

We begin with a brief sampler of some mathematical topics that are useful for reading the later chapters. The description does not aim at being rigorous nor comprehensive. Rather, the purpose is to allow the reader to quickly update his and her knowledge and also it serves the purpose of establishing the notation used in this book.

1.1
Complex Numbers

For the mathematical description of oscillations and waves, the use of complex exponential functions is very practical. For example, a plane wave traveling in x-direction can be represented mathematically by

$$u(x,t) = u_0 \mathrm{e}^{\mathrm{i}(kx-\omega t)}. \tag{1.1}$$

This is the notation we will use in this book. Here, i denotes the imaginary unit defined by $\mathrm{i}^2 = -1$. In engineering, quite often, the letter j is often used instead of i to avoid confusion with the symbol for the electric current. It is also common to write $u(x,t) = u_0 \mathrm{e}^{\mathrm{j}(\omega t - kx)}$. This has no physical consequence, of course. However, it does make a difference in the mathematical formalism, when the first derivative (or, more general, uneven-order derivatives) occur, as it is the case, for example, in the paraxial wave equation.

A complex number z has a real part, denoted as $\Re(z)$, and an imaginary part, $\Im(z)$,

$$z = a + \mathrm{i}b = \Re(z) + \mathrm{i}\,\Im(z). \tag{1.2}$$

Here, $a = \Re(z)$ and $b = \Im(z)$ are real-valued numbers. Using them like Cartesian coordinates, z is represented graphically by its position in the *complex plane* (Figure 1.1).

For the description of a wave that is a harmonic oscillation in space and time, the use of complex exponential functions using polar coordinates is convenient as in (1.1). The exponential form of a complex number is introduced by *Euler's*

Introduction to Micro- and Nanooptics, First Edition. Edited by J. Jahns, S. Helfert.

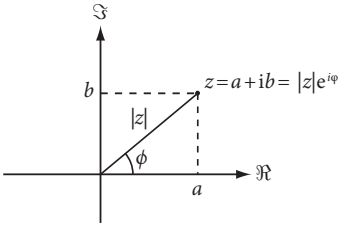


Figure 1.1 Graphical representation of a complex number in the complex plane.

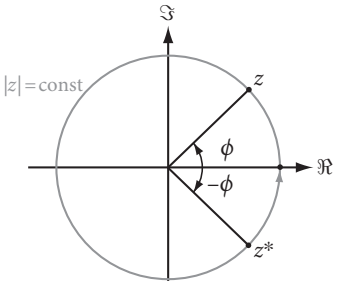


Figure 1.2 Graphical representation of a complex number z using polar coordinates. z^* is the conjugate of z.

equation

$$z = |z|\, \mathrm{e}^{\mathrm{i}\phi} = |z|(\cos\phi + \mathrm{i}\sin\phi)\,. \tag{1.3}$$

Here, $|z|$ is the *modulus* of z with $|z|^2 = a^2 + b^2$. ϕ is called the *argument* or the *phase* of z (Figure 1.2). It is $\phi = \arg(z) = \arctan(b/a)$. In turn, one obtains the Cartesian coordinates from the polar coordinates by $a = |z|\cos\phi$ and $b = |z|\sin\phi$. By varying ϕ, z moves on a circle in the complex plane with a periodicity of 2π. Hence, there is an ambiguity in the polar representation: for a specific point in the complex plane described by the pair of coordinates (a, b), all polar coordinates of the form $(r, \phi + m2\pi)$ with $m = 0, \pm1, \pm2, \ldots$ also represent the same point. This 2π-phase ambiguity is an important aspect of all wave phenomena.

Finally, we introduce the *conjugate* of a complex number. Two numbers z_1 and z_2 are conjugate to each other if their real parts are the same and their imaginary parts differ by a minus sign. The complex conjugate number is denoted either by a bar, $\overline{z}$, or by a star, z^*. Here, we use the latter notation. Thus, we can write

$$z_1 = a_1 + \mathrm{i}b_1 \quad \text{and} \quad z_2 = a_1 - \mathrm{i}b_1 = z_1^*\,. \tag{1.4}$$

Obviously, two conjugate complex numbers z and z^* have the same magnitude. Their geometrical positions are symmetric about the real axis. In exponential notation, the *complex conjugate* of z is $z^* = |z|\mathrm{e}^{-\mathrm{i}\phi}$. In optics, the so-called *intensity* of a wave is of importance, given by the magnitude square of the complex amplitude. The magnitude square of a complex number is given as $|z|^2 = zz^*$. Note that, in general, $|z|^2 \neq z^2$ for a complex number.

1.2 Fourier Transformation

The Fourier transformation is probably the most important mathematical signal transformation. It is widely used for signal analysis, processing, and coding. The most prominent modern application is the encoding of streaming audio and video signals using the MPEG format as well as static images according to the JPEG standard. These formats are based on the *discrete cosine transformation*, a variation of the Fourier transformation suitable for discretized real-valued signals.

The Fourier transformation represents a function $f(x)$ (which we assume to be continuous here) as a linear superposition of sine- and cosine-functions. Using the complex notation of (1.3), in the one-dimensional case, we write

$$f(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \tilde{f}(k) \mathrm{e}^{\mathrm{i}kx} \mathrm{d}k = \mathcal{F}_k^{-1}[\tilde{f}(k)]. \tag{1.5}$$

$\tilde{f}(k)$ is called the Fourier transform (also the *Fourier spectrum*) of $f(x)$, that is,

$$\tilde{f}(k) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x) \mathrm{e}^{-\mathrm{i}kx} \mathrm{d}x = \mathcal{F}_x[f(x)]. \tag{1.6}$$

Here, $1/\sqrt{2\pi}$ is a normalization factor which warrants that

$$\mathcal{F}_k^{-1} \mathcal{F}_x[f(x)] = f(x). \tag{1.7}$$

At this point, let us make a few remarks about the terminology: the term Fourier *transformation* refers to the mathematical operation, whereas the Fourier *transform* denotes the mathematical function. The operator $\mathcal{F}_x$ represents the *forward* Fourier transformation with respect to the x-coordinate, $\mathcal{F}_k^{-1}$ the *inverse* Fourier transformation. We will use the tilde to denote the Fourier transform: $\mathcal{F}_x[f(x)] = \tilde{f}(k)$. k is the Fourier-conjugate variable to x. It is called the *angular frequency* coordinate and relates to the *oscillation frequency* ν_x by $k = 2\pi\nu_x$. If we assume that x is a spatial coordinate, as for a spatial optical wave field, then ν_x has the physical meaning of a *spatial frequency*, that is, its physical unit is an inverse length measured in m^{-1} or $\mu\mathrm{m}^{-1}$, for example. For an optical wave field, the physical meaning of the spatial frequency relates to the angle of propagation relative to a specific coordinate-axis, hence the Fourier transform $\tilde{f}(k)$ is often called the *angular spectrum* of $f(x)$.

As mentioned earlier, it is a matter of definition whether the exponent in (1.5) and (1.6), respectively, is written with a positive or with a negative sign. To be in agreement with the notation as in (1.1) for a single plane wave, the positive sign is used in (1.5). In contrast, when we express the Fourier transformation of a *temporal* signal $g(t)$ with respect to the time coordinate t, we use the negative sign in the Fourier expansion

$$g(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \tilde{g}(\omega) \mathrm{e}^{-\mathrm{i}\omega t} \mathrm{d}\omega = \mathcal{F}_\omega[\tilde{g}(\omega)]. \tag{1.8}$$

In this case, the inverse operation is

$$\tilde{g}(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} g(t) \mathrm{e}^{\mathrm{i}\omega t} \mathrm{d}t = \mathcal{F}_t^{-1}[g(t)]\,. \tag{1.9}$$

Sometimes, it is convenient to express the Fourier expansion in terms of the oscillation frequency ν_t and the spatial frequency ν_x, respectively, rather than the angular frequencies ω and k. In this case, the Fourier expansion of a spatial function $f(x)$ is

$$f(x) = \int_{-\infty}^{\infty} \tilde{f}(\nu_x) \mathrm{e}^{+\mathrm{i}2\pi\nu_x x} \mathrm{d}\nu_x \tag{1.10}$$

and

$$\tilde{f}(\nu_x) = \int_{-\infty}^{\infty} f(x) \mathrm{e}^{-\mathrm{i}2\pi\nu_x x} \mathrm{d}x\,, \tag{1.11}$$

and accordingly for $g(t)$. When compared with (1.5) and (1.6), we note that here the normalization factor is one. This leads to the significant difference in the "DC value" of the signal given by

$$\tilde{f}(\nu_x = 0) = \int_{-\infty}^{\infty} f(x) \mathrm{d}x\,. \tag{1.12}$$

In comparison, when we use (1.6), we see that $\tilde{f}(k = 0)$ yields a different (one might say: "wrong") DC value due to the normalization factor $1/\sqrt{2\pi}$, that is,

$$\tilde{f}(k = 0) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x) \mathrm{d}x\,. \tag{1.13}$$

For further reference, Table 1.1 shows several functions relevant to this text and their Fourier transforms. First, some definitions for

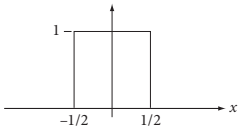
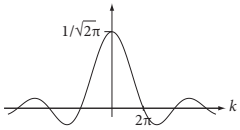
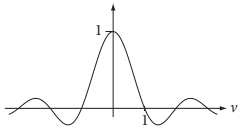
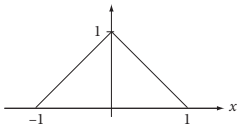
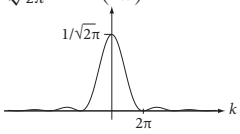
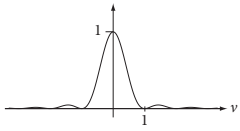
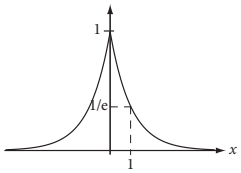
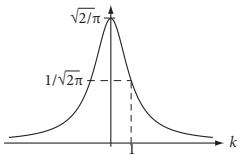
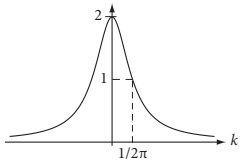
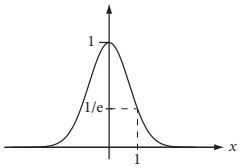
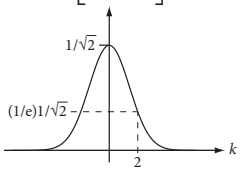
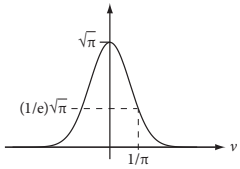
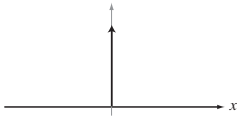
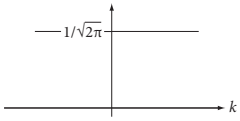
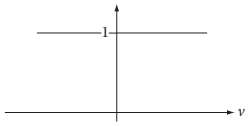
- the *rect-function* $\mathrm{rect}(x)$:

$$\mathrm{rect}(x) = \begin{cases} 1 & \text{if} \quad |x| \le 1/2 \\ 0 & \text{else} \end{cases} \tag{1.14}$$

- the *sinc-function* $\mathrm{sinc}(x)$:

$$\mathrm{sinc}(x) = \frac{\sin(\pi x)}{\pi x} \tag{1.15}$$

Table 1.1 Fourier transformation of rect- and tri-function, exponential, Gaussian and Delta-function.

Function $f(x)$	Fourier transform $\tilde{f}(k)$ (1.6)	Fourier transform $\tilde{f}(\nu)$ (1.11)
$\mathrm{rect}(x)$ 	$\frac{1}{\sqrt{2\pi}}\mathrm{sinc}\left(\frac{k}{2\pi}\right)$ 	$\mathrm{sinc}(\nu)$ 
$\mathrm{tri}(x)$ 	$\frac{1}{\sqrt{2\pi}}\mathrm{sinc}^2\left(\frac{k}{2\pi}\right)$ 	$\mathrm{sinc}^2(\nu)$ 
$\exp(-\lvert x\rvert)$ 	$\sqrt{\frac{2}{\pi}}\frac{1}{1+k^2}$ 	$\frac{2}{1+(2\pi\nu)^2}$ 
$\exp(-x^2)$ 	$\frac{1}{\sqrt{2}}\exp\left[-\left(\frac{k}{2}\right)^2\right]$ 	$\sqrt{\pi}\exp[-(\pi\nu)^2]$ 
$\delta(x)$ 	$\frac{1}{\sqrt{2\pi}}$ 	1 

- the *triangle function* $\mathrm{tri}(x)$:

$$\mathrm{tri}(x) = \begin{cases} 1-|x| & \text{if} \quad |x| \le 1 \\ 0 & \text{else} \end{cases} . \tag{1.16}$$

Remark

Alternative definition of the sinc-function: In the literature, the sinc-function is often defined as

$$\operatorname{sinc}(x) = \frac{\sin(x)}{x}\,, \tag{1.17}$$

that is, without the factor π. Here, however, we will use the definition according to (1.15) since it has the convenient property that the zeros occur at integer values of x. The reader should be aware of the two different definitions.

With the following definitions for the (unnormalized) *Gaussian function*,

$$f(x) = \mathrm{e}^{-x^2}\,, \tag{1.18}$$

and the (unnormalized) *Lorentzian function*,

$$f(x) = \frac{1}{1+x^2}\,, \tag{1.19}$$

keep in mind the following list of Fourier transform pairs shown graphically in Table 1.1 and listed in Table 1.2.

In optics, one usually considers functions that depend on more than one coordinate. The extension of the Fourier transformation to multidimensional functions is straightforward due to its linearity. For example, the angular spectrum of a 2D signal $f(x, y)$ is given as

$$\mathcal{F}_{xy}[f(x,y)] = \tilde{f}(\nu_x, \nu_y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\mathrm{e}^{-\mathrm{i}2\pi(\nu_x x + \nu_y y)}\mathrm{d}x\mathrm{d}y\,. \tag{1.20}$$

We note that for the special situation of a 2D function that is separable in x and y, that is, $f(x, y) = f_1(x)\, f_2(y)$, the transform can be expressed as the product of two 1D transformations, that is,

$$\mathcal{F}_{xy}[f_1(x)\, f_2(y)] = \tilde{f}_1(\nu_x)\, \tilde{f}_2(\nu_y)\,. \tag{1.21}$$

Sometimes, one encounters situations that exhibit radial symmetry. A specific example would be diffraction at a circular aperture. In that case, it may be convenient to carry out the calculations in circular coordinates r and ϕ which are given by

$$x = r\cos\phi \quad \text{and} \quad y = r\sin\phi \tag{1.22}$$

Table 1.2 Fourier transform pairs.

rect	↔	sinc
tri	↔	sinc^2
exponential	↔	Lorentzian
Gaussian	↔	Gaussian
Delta-function	↔	const

with $r^2 = x^2 + y^2$ and $\tan\phi = y/x$. For the spatial frequency domain, one has

$$\nu_x = \rho\cos\theta \quad \text{and} \quad \nu_y = \rho\sin\theta \tag{1.23}$$

with $\rho^2 = \nu_x^2 + \nu_y^2$ and $\tan\theta = \nu_y/\nu_x$. With this, we can express the exponent in (1.20) as

$$\nu_x x + \nu_y y = r\rho(\cos\phi\cos\theta + \sin\phi\sin\theta) = r\rho\cos(\phi - \theta)\,. \tag{1.24}$$

Hence, for the object $u(r,\phi)$, the 2D Fourier transform in radial coordinates is given as

$$\begin{aligned}\tilde{u}(\rho,\theta) &= \int_{\phi=0}^{2\pi}\int_{r=0}^{\infty} u(r,\phi)\mathrm{e}^{-\mathrm{i}2\pi r\rho\cos(\phi-\theta)} r\,\mathrm{d}r\,\mathrm{d}\phi \\ &= \int_{r=0}^{\infty} r\left[\int_{\phi=0}^{2\pi} u(r,\phi)\mathrm{e}^{-\mathrm{i}2\pi r\rho\cos(\phi-\theta)}\mathrm{d}\phi\right]\mathrm{d}r\,.\end{aligned} \tag{1.25}$$

For the special case of circular symmetry, when the object function is independent of ϕ, that is, $u = u(r)$, this becomes

$$\tilde{u}(\rho) = 2\pi\int_0^{\infty} u(r)\,r\,J_0(2\pi\rho r)\mathrm{d}r\,. \tag{1.26}$$

Here, we have used the identity

$$\int_{\phi=0}^{2\pi}\mathrm{e}^{-\mathrm{i}2\pi r\rho\cos(\phi-\theta)}\mathrm{d}\phi = 2\pi J_0(2\pi r\rho)\,. \tag{1.27}$$

Here, $J_0(.)$ is the zeroth Bessel function (of the first kind). The integral transformation in (1.26) is also known as the *Hankel transformation* of the function $u(r)$.

1.2.1 Basic Fourier Rules

In the following, several useful rules for the Fourier transformation are listed which the reader may verify as an exercise. For simplicity, just the one-dimensional case will be considered. Usually, the extension to 2D is straightforward. In order to avoid the normalization factor, we express the Fourier transformation in terms of the spatial frequency variable ν (dropping the index 'x').

Linearity For a function which can be expressed as a linear combination of other functions, the Fourier transform is also given as the linear superposition of the individual transforms, that is,

$$f(x) = ag(x) + bh(x) \rightarrow \tilde{f}(\nu) = a\tilde{g}(\nu) + b\tilde{h}(\nu)\,. \tag{1.28}$$

Scaling If we scale a function in x-direction by a factor with $a > 0$, then its Fourier transform scales with $1/a$, that is,

$$f(x) \rightarrow f(ax) \Rightarrow \tilde{f}(\nu) = \frac{1}{a}\tilde{f}\left(\frac{\nu}{a}\right) \tag{1.29}$$

because with $x' = ax$, we can write

$$\int_{-\infty}^{\infty} f(ax)\mathrm{e}^{\mathrm{i}2\pi\nu x}\mathrm{d}x = \frac{1}{a}\int_{-\infty}^{\infty} f(x')\mathrm{e}^{\mathrm{i}2\pi(\nu/a)x'}\mathrm{d}x' . \tag{1.30}$$

Example

Scaling of a rect-function We calculate the Fourier transform of $f(x) = \mathrm{rect}(ax)$ for $a > 0$. We split up the Fourier integral into its real and imaginary part and make use of the fact that the rect-function is zero for $|x| > 1/2a$:

$$\int_{-\infty}^{\infty} \mathrm{rect}(ax)\mathrm{e}^{-\mathrm{i}2\pi\nu_x x}\mathrm{d}x = \int_{-1/2a}^{1/2a} \cos(2\pi\nu_x x)\mathrm{d}x + \mathrm{i}\int_{-1/2a}^{1/2a} \sin(2\pi\nu_x x)\mathrm{d}x . \tag{1.31}$$

Due to the antisymmetry of the sine-function, the second integral on the right-hand side is zero. Hence, we can continue with the integral over the cosine-function

$$\int_{-1/2a}^{1/2a} \cos(2\pi\nu_x x)\mathrm{d}x = \frac{1}{2\pi\nu_x}\sin(2\pi\nu_x x)\Big|_{-1/2a}^{1/2a} . \tag{1.32}$$

With

$$\sin(2\pi\nu_x x)\Big|_{-1/2a}^{1/2a} = \sin(\pi\nu_x/a) - \sin(-\pi\nu_x/a) = 2\sin(\pi\nu_x/a) \tag{1.33}$$

we obtain

$$\int_{-1/2a}^{1/2a} \cos(2\pi\nu_x x)\mathrm{d}x = \frac{\sin(\pi\nu_x/a)}{\pi\nu_x} = \frac{1}{a}\frac{\sin(\pi\nu_x/a)}{\pi\nu_x/a} , \tag{1.34}$$

and thus we can finally write

$$\int_{-\infty}^{\infty} \mathrm{rect}(ax)\mathrm{e}^{-\mathrm{i}2\pi\nu_x x}\mathrm{d}x = \frac{1}{a}\mathrm{sinc}(\nu_x/a) . \tag{1.35}$$

Figure 1.3 shows the situation for two different scaling factors: $a = 1$ and $a = 2$.

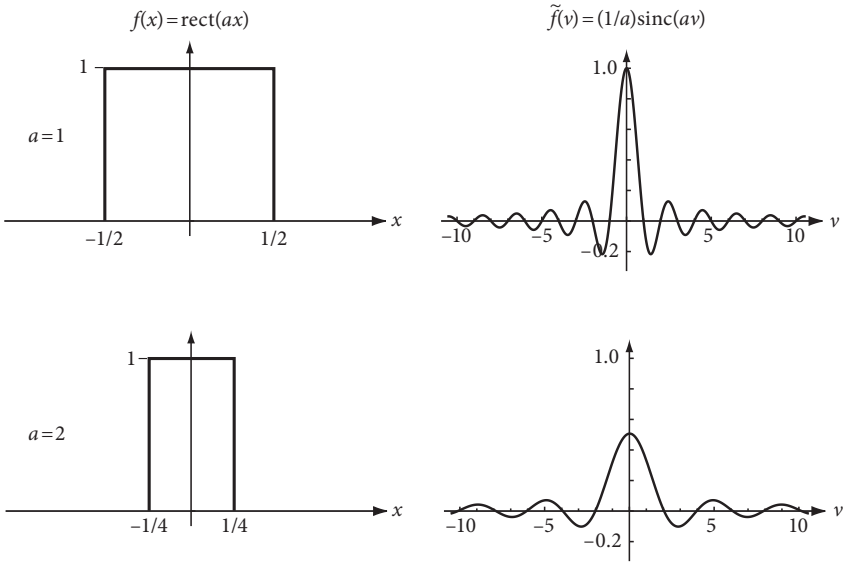


Figure 1.3 Scaled rect-function ($a = 1, 2$) and its Fourier transform (1.33).

Mirror symmetry (even functions) For a symmetric function, the Fourier transform reduces to a *cosine transform*, that is,

$$f(x) = f(-x) \Rightarrow \tilde{f}(\nu) = \int_{-\infty}^{\infty} f(x)\cos(2\pi\nu x)\mathrm{d}x\,. \tag{1.36}$$

Similarly, odd functions with the property $f(x) = -f(-x)$ can be developed into sine-functions.

Hermitian functions A more general statement is: if a function is Hermitian, then its Fourier transform is real-valued, for example,

$$f(x) = f^*(-x) \Rightarrow \tilde{f}(\nu) = \tilde{f}^*(\nu)\,. \tag{1.37}$$

Here, f^* is the complex conjugate of f.

Shift theorem A shift of the function $f(x)$ towards positive x-values by a distance s leads to a phase factor with a negative sign, for example,

$$f(x) \to f(x - x_s) \Rightarrow \tilde{f}(\nu) \to \tilde{f}(\nu)\mathrm{e}^{-\mathrm{i}2\pi\nu x_s}\,. \tag{1.38}$$

The shift theorem can be very useful in conjunction with the convolution theorem, as we will see in the later example.

Translational symmetry (periodic functions) In many areas, periodic functions play an important role, for example, in diffractive optics. We denote the period in x-direction with the letter p in order to write $f(x) = f(x + p)$. In this case, the integral in (1.5) can be replaced by a discrete sum and $f(x)$ can thus be expressed as a series of harmonic functions, namely,

$$f(x) = f(x + p) \Rightarrow f(x) = \sum_{-\infty}^{\infty} a_n \mathrm{e}^{\mathrm{i}2\pi nx/p}\,, \tag{1.39}$$

where the Fourier coefficients a_n are given as

$$a_n = \frac{1}{p}\int_0^p f(x)\mathrm{e}^{-\mathrm{i}2\pi nx/p}\mathrm{d}x\,. \tag{1.40}$$

The following properties of the Fourier transformation relate to the situations where a function f is given as the sum or the product of two other functions.

Convolution theorem If a function can be expressed as the product of two functions, then its Fourier transform is given as the convolution of the two respective Fourier transforms, that is,

$$f(x) = g(x)h(x) \Rightarrow \tilde{f}(\nu) = \int_{-\infty}^{\infty} \tilde{g}(\nu')\tilde{h}(\nu - \nu')\mathrm{d}\nu' = \tilde{g}(\nu) * \tilde{h}(\nu)\,. \tag{1.41}$$

Here, the star $*$ denotes the convolution operation. This theorem can also be applied to the inverse case: if a function can be expressed as the convolution of two functions, then its Fourier spectrum is given as the product of the two respective Fourier transform spectra

$$f(x) = g(x) * h(x) \Rightarrow \tilde{f}(\nu) = \tilde{g}(\nu)\tilde{h}(\nu)\,. \tag{1.42}$$

Autocorrelation function A special case of the rule above occurs when a function f is given when $h(x) = g^*(x)$. In this case, the Fourier transform is the autocorrelation of a function of $\tilde{g}$, that is,

$$f(x) = |g(x)|^2 \Rightarrow \tilde{f}(\nu) = \int_{-\infty}^{\infty} \tilde{g}(\nu')\tilde{g}^*(\nu' - \nu)\mathrm{d}\nu'\,. \tag{1.43}$$

This rather simple statement represents an important physical theorem, the *Wiener–Khinchin theorem*. However, we will not discuss its general significance here.

Parseval's theorem (also known as Plancherel's theorem) The energy of a signal in the x-domain is equal to the energy of its Fourier transform in the ν-domain, that

is,

$$\int_{-\infty}^{\infty} |f(x)|^2 \mathrm{d}x = \int_{-\infty}^{\infty} |\tilde{f}(\nu)|^2 \mathrm{d}\nu \,. \tag{1.44}$$

If instead of the frequency ν the angular frequency k is used, a normalization factor $1/2\pi$ comes in and

$$\int_{-\infty}^{\infty} |f(x)|^2 \mathrm{d}x = \frac{1}{2\pi} \int_{-\infty}^{\infty} |\tilde{f}(k)|^2 \mathrm{d}k \,. \tag{1.45}$$

Fourier transform of the derivative function

$$f(x) = \frac{\mathrm{d}g(x)}{\mathrm{d}x} \Rightarrow \tilde{f}(\nu) = (\mathrm{i}2\pi\nu)\tilde{g}(\nu) \tag{1.46}$$

This can be derived directly from (1.10). For a derivative in the Fourier domain, one gets

$$\tilde{f}(\nu) = \frac{\tilde{g}(\nu)}{\mathrm{d}\nu} \Rightarrow f(x) = (-\mathrm{i}2\pi x)g(x) \,. \tag{1.47}$$

1.3 Maxwell's Equations

An optical wave is an electromagnetic phenomenon and therefore its propagation and interaction with matter are described, in general, by Maxwell's equations. However, there are different ways of writing Maxwell's equations and it depends on the situation regarding which form is appropriate. In the simplest form, they may be written as

$$\nabla \times \boldsymbol{E} + \frac{\partial \boldsymbol{B}}{\partial t} = 0 \,, \tag{1.48}$$

$$\nabla \cdot \boldsymbol{E} = \frac{\rho}{\varepsilon} \,, \tag{1.49}$$

$$\nabla \times \boldsymbol{B} - \varepsilon\mu \frac{\partial \boldsymbol{E}}{\partial t} = \mu \boldsymbol{j} \,, \tag{1.50}$$

$$\nabla \cdot \boldsymbol{B} = 0 \,. \tag{1.51}$$

Here, $\boldsymbol{E}$ is the *electric field* (or *electric field strength*) and $\boldsymbol{B}$ is the *magnetic flux density* (or sometimes, typically in textbooks on physics, simply *magnetic field*). ε is the *electric permittivity* and μ is the *magnetic permeability*. σ denotes the *electric conductivity*, ρ denotes the *charge density*, and $\boldsymbol{j}$ is the *electric current density*. The symbol ∇ is the nabla operator and $\cdot$ denotes the vector product. In the following, we will use the notation $\dot{\boldsymbol{E}}$ for the first partial derivative with time, $\partial \boldsymbol{E}/\partial t$, and $\ddot{\boldsymbol{E}} = \partial^2 \boldsymbol{E}/\partial t^2$ for the

second partial derivative. Equations (1.49) and (1.50) are referred to as the *inhomogeneous Maxwell's equations* since they contain the electric charge density and the current density, while the other two are called the *homogeneous Maxwell's equations*.

$\boldsymbol{E}$ and $\boldsymbol{B}$ are functions of three spatial coordinates (for example, the Cartesian coordinates x, y and z) and the time coordinate t. In isotropic and homogeneous media, the "material constants" ε, μ, ρ and σ are constant. For simplicity, we assume at the beginning that they do not depend on the fields. In this case, Maxwell's equations are linear. Linearity means that if $\boldsymbol{E}_1$ and $\boldsymbol{E}_2$ are solutions as well as $\boldsymbol{B}_1$ and $\boldsymbol{B}_2$, then all linear combinations $a_1\boldsymbol{E}_1 + a_2\boldsymbol{E}_2$ and $b_1\boldsymbol{B}_1 + b_2\boldsymbol{B}_2$ represent solutions, too. The assumption of linearity is not always justified, in particular, when the fields become very large. The generally nonlinear dependency of ε and μ shows up in describing electrooptic and magnetooptic effects.

The set of equations (1.48)–(1.51) is complemented by the equation that relates the electric current density with the $\boldsymbol{E}$-field, that is,

$$\boldsymbol{j} = \sigma\boldsymbol{E}\,. \tag{1.52}$$

The vast majority of optical media for optical propagation is dielectric, in which case $\sigma = 0$. The influence of metallic coatings as used for mirrors and apertures, for example, is often not considered. It does exist, but for the simplified models used in conventional optics, it can usually be disregarded. However, there are certain phenomena where the influence of metallic layers and structures is important and even exploited. This is the case for certain nanooptical and plasmonic devices.

In this context, we have to consider the other material parameters, ε and μ. Classical optics mostly covers the case where $\varepsilon > 0$ and $\mu \approx 1$ (see Figure 1.4). In the (ε, μ)-diagram, many metals are located on the line $\mu \approx 1$, but for them $\varepsilon < 0$. The

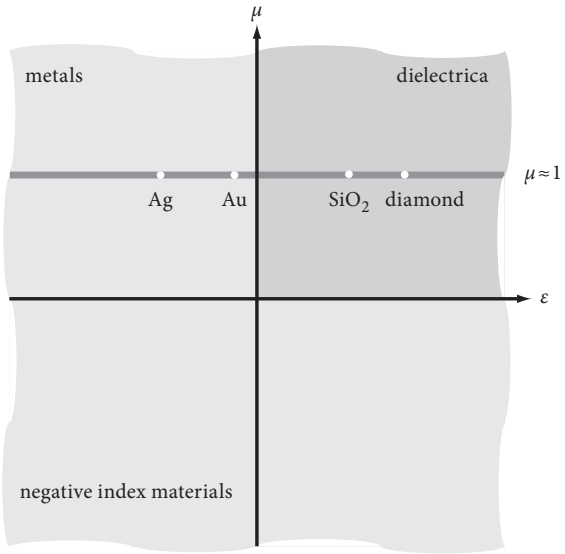


Figure 1.4 Diagram with the material parameters ε and μ as the coordinates.

quadrant on the lower left with $\varepsilon < 0$ and $\mu < 0$ represents the area of *negative-index materials* with very unusual electromagnetic behavior. Such materials are not known in nature (at least, so far), but they can be synthesized by subwavelength-structuring.

The material parameters depend on the molecular structure of a material and/or the geometric structure of a micro- or nanodevice. For most materials and devices, the bulk quantities $\boldsymbol{E}$ and $\boldsymbol{B}$ are sufficient. However, certain aspects suggest that it may sometimes be more convenient to use new field quantities that take the material properties into account. For this purpose, one introduces the *electric displacement density* $\boldsymbol{D}$

$$\boldsymbol{D} = \varepsilon_0 \boldsymbol{E} + \boldsymbol{P} \,. \tag{1.53}$$

$\boldsymbol{P}$ is the *electric polarization (density)*. In a dielectric medium, an electric field causes no current flow, but the induction of dipoles. $\boldsymbol{P}$ is the dipole moment per unit volume. For a linear, homogeneous and isotropic medium, $\boldsymbol{P}$ and $\boldsymbol{E}$ are related by

$$\boldsymbol{P} = \varepsilon_0 \chi \boldsymbol{E} = \varepsilon_0 (\varepsilon_\mathrm{r} - 1) \boldsymbol{E} \,. \tag{1.54}$$

Here, χ is the *electric susceptibility* and ε_r is the relative permittivity. By combining (1.53) and (1.54), one obtains

$$\boldsymbol{D} = \varepsilon_0 \varepsilon_\mathrm{r} \boldsymbol{E} = \varepsilon \boldsymbol{E} \,. \tag{1.55}$$

Simplified, one may say that $\boldsymbol{D}$ is the $\boldsymbol{E}$-field in a medium with the materials properties taken into account. Expressed by $\boldsymbol{D}$ rather than $\boldsymbol{E}$, Maxwell's equation reads as

$$\nabla \cdot \boldsymbol{D} = \rho \,. \tag{1.56}$$

The analogous expressions for the *magnetic field (strength)* $\boldsymbol{H}$ are

$$\boldsymbol{H} = \frac{1}{\mu_0} \boldsymbol{B} - \boldsymbol{M} \,, \tag{1.57}$$

$$\boldsymbol{B} = \mu_0 \mu_\mathrm{r} \boldsymbol{H} \,, \tag{1.58}$$

$$\nabla \times \boldsymbol{H} - \frac{\partial \boldsymbol{D}}{\partial t} = \boldsymbol{j} \,. \tag{1.59}$$

$\boldsymbol{M}$ is the *magnetic polarization* and μ_r is the *relative magnetic permeability*.

1.4 Boundary Conditions

As mentioned, structuring of a medium, in particular, at the subwavelength scale, allows one to "engineer" the optical parameters. We shall learn about this topic

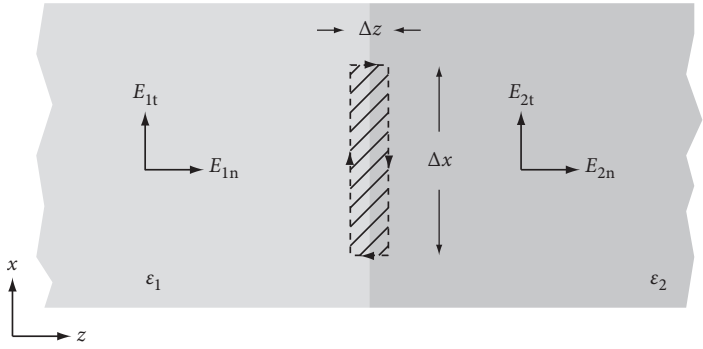


Figure 1.5 Path of integration along the boundary of two media.

in the later sections of this book. We prepare these issues by looking first at the boundary conditions for the components of the electric and magnetic field. For this purpose, we consider the interface between two media which differ in the values of the electric permittivity ε (Figure 1.5).

For the derivation of the boundary conditions, one may apply the first of Maxwell's equations. In its integral form, it reads

$$\oint \boldsymbol{E} \cdot \mathrm{d}\boldsymbol{r} = -\iint \dot{\boldsymbol{B}} \cdot \mathrm{d}\boldsymbol{a}\,. \tag{1.60}$$

The integral on the left-hand side is a line integral along the closed path indicated in the figure by the dotted line. The integral on the right-hand side sums up across the hatched area which is enclosed by the path. $\mathrm{d}\boldsymbol{r}$ is the path differential, and $\mathrm{d}\boldsymbol{a}$ is the surface differential normal to the surface. $\dot{\boldsymbol{B}}$ is the time derivative of $\boldsymbol{B}$. By applying (1.60) to the situation of Figure 1.5, for the left-hand side, one obtains

$$\begin{aligned}\oint \boldsymbol{E} \cdot \mathrm{d}\boldsymbol{r} &\approx \Delta x(E_{1\mathrm{t}} - E_{2\mathrm{t}}) + \frac{\Delta z}{2}[(E_{1\mathrm{n}} + E_{2\mathrm{n}}) - (E_{2\mathrm{n}} + E_{1\mathrm{n}})] \\ &= \Delta x(E_{1\mathrm{t}} - E_{2\mathrm{t}})\,.\end{aligned} \tag{1.61}$$

Here, $E_{k\mathrm{n}}$ and $E_{k\mathrm{t}}$ are the normal and tangential components of the $\boldsymbol{E}$-field in both media with $k = 1, 2$. In (1.61), we assume that the normal and tangential components are approximately the same at the upper and lower sections of the integration path.

The right-hand side of (1.60) can be evaluated as

$$\iint \dot{\boldsymbol{B}} \cdot \mathrm{d}\boldsymbol{a} \approx \Delta x \Delta z \langle \dot{\boldsymbol{B}} \rangle\,. \tag{1.62}$$

Here, $\langle \dot{\boldsymbol{B}} \rangle$ is the average value of the derivative of $\boldsymbol{B}$. Now, we assume that we decrease $\Delta z \to 0$. For a finite value of $\dot{\boldsymbol{B}}$, the integral in (1.62) will go to zero so that we obtain

$$\oint \boldsymbol{E} \cdot \mathrm{d}\boldsymbol{r} \approx \Delta x(E_{1\mathrm{t}} - E_{2\mathrm{t}}) \approx 0\,, \tag{1.63}$$

and hence

$$E_{1t} = E_{2t}\,. \tag{1.64}$$

This means that the tangential component of the electric field is continuous at a boundary. This is not true, however, for the $\boldsymbol{D}$-field. The tangential component D_t is discontinuous due to the surface charge at the interface. With the same arguments as above, one can show that for a linear, isotropic medium,

$$\frac{D_{1t}}{\varepsilon_1} = \frac{D_{2t}}{\varepsilon_2}\,. \tag{1.65}$$

Another important result relates to the normal components. From the second of Maxwell's equations ($\nabla \cdot \boldsymbol{D} = \rho$), one can derive the behavior of the normal component of the $\boldsymbol{D}$-field. We do not enter into a detailed discussion, but rather give the basic arguments. In its integral form, one can express the second of Maxwell's equations as

$$\iint \boldsymbol{D} \cdot \mathrm{d}\boldsymbol{a} = \int \rho \mathrm{d}V = q\,. \tag{1.66}$$

Here, the integration takes place over the surface indicated as a cross-section by the dotted line in Figure 1.5. q is the electrical charge contained in the integration volume. For decreasing dimensions of the integration volume (and surface, respectively), q is approximated by the surface charge σ at the interface. One can then argue that at the interface of two dielectric media, the extension of surface charges into the media is so small that for $\Delta z \to 0$, the amount of surface charge between two dielectric media $\sigma \to 0$ from which

$$D_{1n} = D_{2n}\,. \tag{1.67}$$

That is, the normal component of the $\boldsymbol{D}$-field, D_n, is continuous at an interface if surface charges can be neglected. The latter assumption is justified in the case of dielectric media. This result will be used later when we discuss the optical properties of microstructured media.

1.4.1 Method of Stationary Phase

The method of stationary phase allows one the approximate calculation of an integral given as

$$\int g(x)\mathrm{e}^{\mathrm{i}\phi(x)}\mathrm{d}x\,. \tag{1.68}$$

We assume that $g(x)$ is a slowly varying function, while $\phi(x)$ oscillates rapidly. We will encounter such integrals, for example, when discussing paraxial diffraction theory. The integral can be estimated if $\phi(x)$ is stationary at certain points, meaning that $\mathrm{d}\phi/\mathrm{d}x = 0$ for some coordinate $x = x_0$. As we will see in the following

calculation, due to the rapid oscillation of $\phi(x)$, significant contributions to the integral only come from that point. For now, let us assume that this is the case at one coordinate. (Remark: If ϕ is stationary at more than one points, one subdivides the axis into several intervals to perform the following calculation.) First, we can evaluate $\phi(x)$ into a Taylor series around x_0, that is,

$$\phi(x) = \phi_0 + \frac{1}{2}\phi_0''(x - x_0)^2 + \dots \tag{1.69}$$

where $\phi_0 = \phi(x_0)$ and $\phi_0'' = \frac{\mathrm{d}^2\phi(x)}{\mathrm{d}x^2}|_{x_0}$. With this, we may write

$$\int g(x)\mathrm{e}^{\mathrm{i}\phi(x)}\mathrm{d}x \approx g(x_0)\mathrm{e}^{\mathrm{i}\phi_0}\int_{x_u}^{x_0}\mathrm{e}^{(\mathrm{i}/2)\phi_0''(x-x_0)^2}\mathrm{d}x\,. \tag{1.70}$$

We introduce the coordinate transformation $\xi^2 = (1/2)\phi_0''(x - x_0)^2$ to write

$$\int \mathrm{e}^{(\mathrm{i}/2)\phi_0''(x-x_0)^2}\mathrm{d}x = \sqrt{\frac{2}{\phi_0''}}\int \mathrm{e}^{\mathrm{i}\xi^2}\mathrm{d}\xi\,. \tag{1.71}$$

We split up the integral on the right-hand side into its real and imaginary part, namely,

$$\int \mathrm{e}^{\mathrm{i}\xi^2}\mathrm{d}\xi = \int \cos\xi^2\mathrm{d}\xi + \mathrm{i}\int \sin\xi^2\mathrm{d}\xi\,. \tag{1.72}$$

These integrals are known as the *Fresnel integrals*. Their calculation yields the values

$$\int \cos\xi^2\mathrm{d}\xi = \int \sin\xi^2\mathrm{d}\xi = \sqrt{\frac{\pi}{2}}\,. \tag{1.73}$$

From Figure 1.6, we see that the value of the integral is determined from the area around $x = x_0$ (shaded in the figure), while the oscillations cancel each other. We may continue to write

$$\int \mathrm{e}^{\mathrm{i}\xi^2}\mathrm{d}\xi = \sqrt{\frac{\pi}{2}}(1 + \mathrm{i}) = \sqrt{\pi}\mathrm{e}^{\mathrm{i}\pi/4}\,. \tag{1.74}$$

Thus, under the assumptions made earlier, we finally obtain

$$\int g(x)\mathrm{e}^{\mathrm{i}\phi(x)}\mathrm{d}x \approx \sqrt{\frac{2\pi}{\phi_0''}}g(x_0)\mathrm{e}^{\mathrm{i}(\phi_0+\pi/4)}\,. \tag{1.75}$$

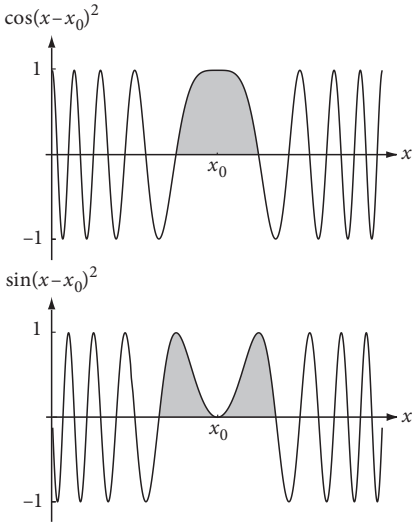


Figure 1.6 Integration of the Fresnel integrals.

Questions

1. What is a complex number?
2. What is the complex conjugate?
3. In the complex plane, what is the location of all complex numbers of the form $e^{i\phi}$?
4. Explain the Fourier transformation in physical terms (a) for a temporal signal and (b) for a spatial signal.
5. What is a spatial frequency?
6. What is the meaning of the shift-theorem of Fourier mathematics?
7. How does the electric field behave at the boundary between two dielectric media?
8. What is the method of stationary phase?

Problems

1. *Complex numbers* For the two complex numbers, $z_1 = |z_1|\, e^{i\phi_1}$ and $z_2 = |z_2|\, e^{i\phi_2}$, calculate (a) the sum, (b) the product, and (c) the quotient.
2. *Multiplication of complex numbers* If z_1 and z_2 are complex numbers, prove that $|z_1 z_2| = |z_1||z_2|$.
3. *Power of complex numbers* Evaluate the expression $\sum_{n=0}^{4} i^n$.
4. *Derivative of complex expression* For $u(x,t) = e^{i(kx-\omega t)}$, calculate $\partial u/\partial x$ and $\partial u/\partial t$.
5. *Real and imaginary part* For $u(x) = e^{iax^2}$ (a real and $a > 0$), draw (or plot) $\Re(u)$ and $\Im(u)$.

6. *Fourier transformation of time signal* Calculate the temporal Fourier transform of the function $u(t) = \mathrm{e}^{\mathrm{i}\nu_0 t} + \mathrm{e}^{-\mathrm{i}\nu_0 t}$.
7. *Fourier transformation of a derivative function* Derive (1.46) from (1.10).
8. *Normalization factor of the Fourier transformation* The normalization factor has to warrant that successive application of forward and backward transformation leads to the original function, that is,

$$\mathcal{F}_k^{-1}\mathcal{F}_x[f(x)] = f(x)\,. \tag{1.76}$$

Show that

$$\int_{-\infty}^{\infty}\left[\int_{-\infty}^{\infty} f(x)\mathrm{e}^{-\mathrm{i}kx}\mathrm{d}x\right]\mathrm{e}^{\mathrm{i}kx'}\mathrm{d}k = 2\pi f(x') \tag{1.77}$$

while

$$\int_{-\infty}^{\infty}\left[\int_{-\infty}^{\infty} f(x)\mathrm{e}^{-\mathrm{i}2\pi\nu x}\mathrm{d}x\right]\mathrm{e}^{\mathrm{i}2\pi\nu x'}\mathrm{d}\nu = f(x')\,. \tag{1.78}$$

Further Reading

1 Jackson, J.D. (1998) *Classical Electrodynamics*, 3rd edn, John Wiley & Sons (Asia) Pte Ltd.

2 James, J.F. (2011) *A Student's Guide to Fourier Transforms*, 3rd edn, Cambridge University Press.