VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ
BRNO UNIVERSITY OF TECHNOLOGY

FAKULTA STROJNÍHO INŽENÝRSTVÍ
ÚSTAV MATEMATIKY
FACULTY OF MECHANICAL ENGINEERING
INSTITUTE OF MATHEMATICS

PHASE-CORRELATION BASED IMAGE REGISTRATION

REGISTRACE OBRAZŮ POMOCÍ FÁZOVÉ KORELACE

DIPLOMOVÁ PRÁCE
MASTER'S THESIS

AUTOR PRÁCE
AUTHOR
Bc. HANA DRUCKMÜLLEROVÁ

VEDOUCÍ PRÁCE
SUPERVISOR
Mgr. JANA PROCHÁZKOVÁ, Ph.D.

BRNO 2010

Abstrakt
Tato práce se zabývá použitím fázové korelace k určení vzájemné rotace, změny měřítka a posunu mezi digitálními obrazy. Fázová korelace je založena na Fourierově transformaci, proto je popsána Fourierova transformace funkcí definovaných na $\mathbb{R}^2$ i diskrétní Fourierova transformace funkcí definovaných na konečném počtu bodů $\{0, 1, \ldots, N-1\}^2, N \in \mathbb{N}$. Dále je pozornost věnována modifikacím fázové korelace, díky nimž metoda umožňuje nalezení parametrů podobnostní transformace i mezi obrazy, které mají vysoký dynamický rozsah a slabě patrné struktury, obsahují aditivní nebo impulzní šum a jsou pořízeny pomocí různých snímačů a optických soustav. Obsahem práce jsou i modifikace metody pro snímky sluneční koróny pořízené během úplných zatmění Slunce, což patří mezi nejobtížnější úlohy registrace obrazů.

Summary
This thesis deals with the use of phase correlation for estimation of mutual rotation, scale-change and translation between digital images. Phase correlation is based on the Fourier transform, therefore the thesis describes the Fourier transform of functions defined on $\mathbb{R}^2$ and the discrete Fourier transform of functions defined on a finite number of points $\{0, 1, \ldots, N-1\}^2, N \in \mathbb{N}$. The thesis describes modifications of the phase correlations for estimation of similarity transformation parameters between images with high dynamic range and faint structures, containing additive or impulse noise and images which are taken with different sensors and optical systems. The thesis also focuses on the modifications of the method used for registration of solar corona images taken during total solar eclipses, which counts among the hardest registration tasks.

klíčová slova
registrace obrazů, Fourierova transformace, fázová korelace, sluneční koróna

key words
image registration, Fourier transform, phase correlation, solar corona

DRUCKMÜLLEROVÁ, H. *Phase-correlation based image registration*, Brno: Brno University of Technology, Faculty of Mechanical Engineering, 2010. (100 pages). Supervisor Mgr. Jana Procházková, PhD.

I declare that I have written the master's thesis *Phase-correlation based image registration* on my own according to advice of my master's thesis supervisor Mgr. Jana Procházková, PhD., and using the sources listed in references.

May 19, 2010 Hana Druckmüllerová

I would like to express thanks to my master's thesis supervisor Mgr. Jana Procházková, PhD. for numerous comments and valuable suggestions on improving my thesis.

Hana Druckmüllerová

Contents

1 Introduction **3**

2 The Fourier transform **5**
2.1 Basic notions 5
2.2 The Dirac distribution 8
2.3 The Fourier transform and the inverse Fourier transform 12
2.3.1 Functions in $\mathcal{L}(\mathbb{R})$ 12
2.3.2 Functions in $\mathcal{L}(\mathbb{R}^2)$ 18
2.4 Properties of the Fourier transform 21
2.5 Convolution 31
2.6 Cross correlation, phase correlation 34

3 The discrete Fourier transform **37**
3.1 The discrete Fourier transform and the inverse discrete Fourier transform . 37
3.2 Properties of the discrete Fourier transform 40
3.3 Discrete periodic convolution 45
3.4 Cross correlation, phase correlation 46

4 Image registration **50**
4.1 Digital image 50
4.2 Registration of idealized images 52
4.2.1 Registration of identical images 52
4.2.2 Registration of shifted images 55
4.2.3 Registration of similar images 58
4.3 Registration of real images 63
4.3.1 Registration of images with respect to their modulation transfer function 65
4.3.2 Registration of images containing additive noise and variable impulse noise, with vignetting and diffuse light 66
4.3.3 Registration of images containing image defects, constant impulse noise and dust 68
4.3.4 Registration of shifted images step by step 71
4.3.5 Registration of similar images step by step 72
4.3.6 Registration of motion-blurred images 73
4.4 Subpixel precision registration 74
4.5 Checking the results 75
4.6 Registration of solar corona images 75
4.6.1 Tangential high-pass filter T_ς 76

4.6.2 Methods and parameters used for solar corona images 78
4.6.3 Registration of images at different wavelengths 79
4.7 Testing registration precision 80
4.7.1 Testing on simulated data 80
4.7.2 Testing on real data 82

5 Conclusion **84**

References **86**

Used symbols **90**

A Proof of equality $\int_0^\infty \frac{\sin x}{x}\,\mathrm{d}x = \frac{\pi}{2}$ **92**
A.1 Integration-theory proof 92
A.2 Fourier-series proof 94

B Proof of equality $\int_0^\infty \frac{|\sin x|}{x}\,\mathrm{d}x = \infty$ **96**

C Proof of equality $\int_0^\infty \frac{\sin^2 x}{x}\,\mathrm{d}x = \infty$ **97**

D Definitions of quadruple improper integral and $\mathcal{L}\,(\mathbb{R}^4)$ **99**

Chapter 1

Introduction

Registration of images generally means finding the mutual geometrical transformations between the images. It is a very important part of image processing since we often need to compose more images together. They are either overlapping by great parts and the purpose of the composition is to produce a better image (lower additive noise etc.) or they are less overlapping and the purpose of the composition is to produce a mosaic of the images. The methods available for image registration depend on whether the images are shifted only, similar (rotated, scaled, shifted) or distorted in a more complicated way.

Various methods for image registration are described in [49]. Some methods are based directly on pixel values, e.g. the correlation methods [13] and the mutual-information method [36]. Such methods are unusable if the images to be registered are exposed differently. Other methods are based on matching of corresponding structures in the images. However, these methods require that these structures are visible and clearly defined in the images. This requirement is usually fulfilled when we register images of a cell structure, other medical specimen or for instance aerial images. Registration of solar corona images is different. The only clearly visible and defined structures in the images are (apart from the Moon edge, which moves in front of the Sun) prominences. They are visible on shorter exposures only. On images taken with longer exposure time, prominences are saturated as well as the neighboring corona. A very contrasty structure is the edge of the saturated part. Structures in the corona are very faint and it is impossible to find matching structures in two images which are so clearly defined that they could be used for image registration.

This problem can be overcome by registration methods which use the image as a whole rather than separate structures in the images. Phase correlation is a method which works with Fourier spectra of the images to be registered and is described e.g. in [13]. The principles for phase correlation for estimating scale change, rotation and shift of images is described in [45]. Its modifications for registration of solar corona images are described in [23]. The methods assume that there is no other geometrical transformation between the images than scale-change, rotation and translation. Other geometrical transformations are usually negligible, especially in images taken with the same equipment.

When implemented in computer programs, phase correlation uses the Fast Fourier transform, which is a fast implementation of the discrete Fourier transform. However, the principles of the method (especially for scaled and rotated images) can be only described by means of the Fourier transform of functions defined on $\mathbb{R}^2$. The Fourier transform, its properties and the phase correlation of functions defined on $\mathbb{R}^2$ is discussed in Chapter 2. Chapter 3 describes the discrete Fourier transform, its properties and the phase correla-

tion of functions defined on a finite number of points arranged in a square. Chapter 4 deals with the phase-correlation based method for image registration and its modifications for less similar images and for images of the solar corona images. It describes the mathematical background and practical aspects of image registration using phase correlation in a detailed way which cannot be found in literature.

Chapter 2

The Fourier transform

2.1 Basic notions

Definition 2.1. Let $f(x)$ be a a complex function of a real variable, i.e. $f : M \to \mathbb{C}$, $M \subseteq \mathbb{R}$,

$$f(x) = u(x) + \mathrm{i}v(x),$$

where $u, v : M \to \mathbb{R}$ are real functions of a real variable. Let $a, b \in \mathbb{R}, a < b$ and let u, v be Reimann integrable [38] on interval $\langle a, b\rangle$. Then we say that function f is *integrable* on $\langle a, b\rangle$ and set [18]

$$\int_a^b f(x)\,\mathrm{d}x = \int_a^b u(x)\,\mathrm{d}x + \mathrm{i}\int_a^b v(x)\,\mathrm{d}x.$$

Analogically, we define the antiderivative of f as

$$\int f(x)\,\mathrm{d}x = \int u(x)\,\mathrm{d}x + \mathrm{i}\int v(x)\,\mathrm{d}x.$$

Remark 2.2. The previous definition shows that the rules for computing with integrals of complex functions of a real variable are the same as those for computing with integrals of real functions of a real variable.

Definition 2.3. Let function $f : \mathbb{R} \to \mathbb{R}$ be Riemann integrable [38] on every interval $\langle a, b\rangle$, $a, b \in \mathbb{R}$, $a < b$. If the following limit exists and is finite

$$\lim_{b\to\infty} \int_a^b f(x)\,\mathrm{d}x = A,$$

we say that integral

$$\int_a^\infty f(x)\,\mathrm{d}x$$

exists (converges, is convergent) [38] and write

$$\int_a^\infty f(x)\,\mathrm{d}x = A.$$

If $A = -\infty$ or $A = \infty$, we say that the integral *diverges* and set

$$\int_a^\infty f(x)\,\mathrm{d}x = A.$$

In a similar way we treat the limit

$$\lim_{a\to-\infty} \int_a^b f(x)\,\mathrm{d}x = B.$$

If both integrals

$$\int_{-\infty}^0 f(x)\,\mathrm{d}x, \quad \int_0^\infty f(x)\,\mathrm{d}x$$

converge, we say that integral

$$\int_{-\infty}^\infty f(x)\,\mathrm{d}x$$

exists (converges) [38] and it is equal to the sum of both integrals

$$\int_{-\infty}^\infty f(x)\,\mathrm{d}x = \int_{-\infty}^0 f(x)\,\mathrm{d}x + \int_0^\infty f(x)\,\mathrm{d}x.$$

Furthermore, if both integrals

$$\int_{-\infty}^0 f(x)\,\mathrm{d}x, \quad \int_0^\infty f(x)\,\mathrm{d}x$$

diverge and have the same sign, we say that integral

$$\int_{-\infty}^\infty f(x)\,\mathrm{d}x$$

diverges and set

$$\int_{-\infty}^\infty f(x)\,\mathrm{d}x = \int_{-\infty}^0 f(x)\,\mathrm{d}x + \int_0^\infty f(x)\,\mathrm{d}x.$$

Definition 2.4 ($\mathcal{L}(\mathbb{R})$). Let us denote $\mathcal{L}(\mathbb{R})$ the space of all functions $\mathbb{R} \to \mathbb{C}$ such that

$$\int_{-\infty}^\infty |f(x)|\,\mathrm{d}x$$

exists and is finite.

Remark 2.5. The $\mathcal{L}(\mathbb{R})$ space has analogical properties with the $\mathcal{L}^1(\mathbb{R})$ space [33]. The $\mathcal{L}^1(\mathbb{R})$ space is defined by means of the Lebesgue integral. However, some of the further proves are based on Riemann integration. We will need integrals which are not defined as Lebesgue integrals, but they converge as Riemann integrals. In image registration, we will work with functions whose values are known in a finite number of points only. Then we can make these functions both Riemann and Lebesgue integrable. Furthermore, the space is usually defined for real functions ($\mathbb{R} \to \mathbb{R}$). In Fourier analysis, we work with complex functions ($\mathbb{R} \to \mathbb{C}$) and Definition 2.1 makes this notion reasonable.

In accordance with Definition 2.3 of the improper integral of a real function of one real variable, we define the improper double integral, which can be easily generalized to more dimensions (the definition for the four-dimensional case is in Appendix D on page 99) and to complex functions of two real variables.

Definition 2.6. Let $f(x, y)$ be a function $\mathbb{R}^2 \to \mathbb{R}$. We write [48]

$$\lim_{(x,y)\to(\infty,\infty)} f(x, y) = L$$

if and only if for every $\varepsilon > 0$ there exists $r \in \mathbb{R}$ such that for every $a > r, b > r$: $|f(a, b) - L| < \varepsilon$. Analogically, such limit is defined for $(x, y) \to (\infty, -\infty)$, $(x, y) \to (-\infty, \infty)$ and $(x, y) \to (-\infty, -\infty)$.

Definition 2.7 (Improper double integral). Let $f(x, y)$ be a function $\mathbb{R}^2 \to \mathbb{R}$. Let $R = \langle a, \infty) \times \langle c, \infty)$, $a, c \in \mathbb{R}$. If the following limits exist and are equal

$$\lim_{(b,d)\to(\infty,\infty)} \int_a^b \left(\int_c^d f(x, y)\,\mathrm{d}y \right) \mathrm{d}x = \lim_{(b,d)\to(\infty,\infty)} \int_c^d \left(\int_a^b f(x, y)\,\mathrm{d}x \right) \mathrm{d}y = A,$$

then we define

$$\iint_R f(x, y)\,\mathrm{d}x\,\mathrm{d}y = A.$$

Analogically, the integral is defined for $R = (-\infty, b\rangle \times \langle c, \infty)$, $R = \langle a, \infty) \times (-\infty, d\rangle$ and $R = (-\infty, b\rangle \times (-\infty, d\rangle$. Furthermore, if all the following integrals exist and are finite, or, in case some of the integrals are infinite, they have the same sign

$$\begin{aligned}
\iint_{\langle 0,\infty)^2} f(x, y)\,\mathrm{d}x\,\mathrm{d}y &= B \\
\iint_{(-\infty,0\rangle\times\langle 0,\infty)} f(x, y)\,\mathrm{d}x\,\mathrm{d}y &= C \\
\iint_{\langle 0,\infty)\times(-\infty,0\rangle} f(x, y)\,\mathrm{d}x\,\mathrm{d}y &= D \\
\iint_{(-\infty,0\rangle^2} f(x, y)\,\mathrm{d}x\,\mathrm{d}y &= E,
\end{aligned}$$

we define

$$\iint_{(-\infty,\infty)^2} f(x, y)\,\mathrm{d}x\,\mathrm{d}y = B + C + D + E.$$

If function $f(x,y)$ is $\mathbb{R}^2 \to \mathbb{C}$, from the definition we integrate the real and imaginary part separately and Definition 2.1 (page 5) is used for each of them.

Definition 2.8 ($\mathcal{L}(\mathbb{R}^2)$). Let us denote $\mathcal{L}(\mathbb{R}^2)$ the space of functions $\mathbb{R}^2 \to \mathbb{C}$ such that

$$\iint_{\mathbb{R}^2} |f(x,y)| \,\mathrm{d}x \,\mathrm{d}y$$

exists and is finite.

Convention 2.9. Let $a,b,c,d \in \mathbb{R} \cup \{-\infty,\infty\}$, $a<b, c<d$ and let $f(x,y) \in \mathcal{L}(\mathbb{R}^2)$. Analogically with Fubini's theorem [38] the double integral of function f on cartesian rectangle $\langle a,b\rangle \times \langle c,d\rangle$ can equivalently be written in the following forms

$$\iint_{\langle a,b\rangle\times\langle c,d\rangle} f(x,y)\,\mathrm{d}x\,\mathrm{d}y = \int_a^b \left(\int_c^d f(x,y)\,\mathrm{d}y\right)\mathrm{d}x = \int_a^b\int_c^d f(x,y)\,\mathrm{d}x\,\mathrm{d}y.$$

Definition 2.10 (Characteristic function of a set). Let X be a set and A its subset. The *characteristic function* of set A relative to set X, χ_A is defined for every $x \in X$ as

$$\chi_A(x) = \begin{cases} 1 & \text{if } x \in A \\ 0 & \text{if } x \notin A. \end{cases}$$

Sometimes, the name *indicator function* is used [9].

2.2 The Dirac distribution

Definition 2.11 (Finite function). A function $f(x,y) : \mathbb{R}^2 \to \mathbb{R}$ is called *finite* [33] if it is equal to zero outside of a cartesian rectangle $\langle a,b\rangle \times \langle c,d\rangle$, where $a,b,c,d \in \mathbb{R}$, $a<b, c<d$.

Remark 2.12. It is possible to define the finite function as a function which is equal to zero outside of a bounded domain (the concept of domain to be defined in [18]). However, it is not necessary here. For every function which is zero outside of a bounded domain, there is a cartesian product of intervals outside of which it is equal to zero. And vice versa, if a function is equal to zero outside of a cartesian product of intervalus $\langle a,b\rangle \times \langle c,d\rangle$, it is equal to zero outside of every domain M such that $\langle a,b\rangle \times \langle c,d\rangle \subset M$.

Theorem 2.13. Let K be the set of all finite functions φ which have continuous derivatives of all orders. Then $\mathcal{K} = (K,+,\cdot)$ is a vector space with standardly defined operations of functions addition and multiplication by a constant.

Proof. It is easy to see that for any functions $\varphi,\psi \in K$ and constants $a,b \in \mathbb{R}$ it holds that $\varphi+\psi \in K$, the neutral element of the additive group $(K,+)$ is a function which is equal to zero everywhere, the inverse element of function φ is function $-\varphi$, addition of functions is commutative, $k\cdot(\varphi+\psi) = k\cdot\varphi + k\cdot\psi$, $(k+l)\cdot\varphi = k\cdot\varphi + l\cdot\varphi$, $(k\cdot l)\cdot\varphi = k\cdot(l\cdot\varphi)$ and $1\cdot\varphi = \varphi$. □

Definition 2.14 (Basic space, basic functions). Space $\mathcal{K}$ from Theorem 2.13 is called the *basic space* and its elements are called *basic functions* [33].

Definition 2.15 (Convergence in $\mathcal{K}$). A sequence $\{\varphi_n\} \subset K$ is called *convergent* to function $\varphi \in K$ [33] if

1. there is a cartesian rectangle $\langle a, b\rangle \times \langle c, d\rangle$, $a, b, c, d \in \mathbb{R}$, $a < b, c < d$, out of which all functions φ_n are equal to zero,
2. the sequences of derivatives $\left\{\frac{\partial^k \varphi_n}{\partial x_1 x_2 \ldots x_k}\right\}$, $k = 0, 1, 2, \ldots$, $x_j \in \{x, y\}$ converge uniformly to derivatives $\frac{\partial^k \varphi}{\partial x_1 x_2 \ldots x_k}$.

Definition 2.16 (Continuousness of linear functionals on $\mathcal{K}$). A linear functional T on the basic space $\mathcal{K}$ is called *continuous* [33] if for every sequence of functions $\{\varphi_n\} \subset K$ which converges to function $\varphi \in K$ it holds $T(\varphi_n) \to T(\varphi)$.

Definition 2.17 (Distribution on $\mathbb{R}^2$). A continuous linear functional T on the basic space $\mathcal{K}$ is called a *distribution* [33] on $\mathbb{R}^2$.

Remark 2.18. According to the previous definition, each function f which is integrable on every cartesian rectangle $\langle a, b\rangle \times \langle c, d\rangle$ generates a distribution since $T_f(\varphi)$ defined by the following formula

$$T_f(\varphi) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y)\varphi(x, y)\,\mathrm{d}x\,\mathrm{d}y$$

is a continuous linear functional defined on $\mathcal{K}$.

Definition 2.19 (Dirac distribution). The *Dirac distribution* [33] is a functional

$$T(\varphi) = \varphi(0, 0)$$

on the basic space $\mathcal{K}$.

The Dirac distribution is usually written as [33]

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \delta(x, y)\varphi(x, y)\,\mathrm{d}x\,\mathrm{d}y,$$

where the symbol $\delta(x, y)$ represents a "function" equal to zero for all $(x, y) \neq (0, 0)$ such that

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \delta(x, y)\,\mathrm{d}x\,\mathrm{d}y = 1.$$

It is obvious that such "function" is even.

Definition 2.19 is the original definition for the δ-distribution introduced by Dirac. It enables us to think of the δ-distribution as an infinitely high, infinitely short impulse. However, there is another definition which is sometimes used in Fourier analysis. Let us also introduce the definition of a δ-distribution used in [34] which is not equivalent with the Dirac's definition. A similar approach is also used in [44].

Definition 2.20 (δ-distribution). The *δ-distribution* [34] is the limit of a sequence of functions $\delta_p(x, y)$, $p \in \mathbb{N}$ such that

$$\lim_{p\to\infty} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \delta_p(x, y)\,\mathrm{d}x\,\mathrm{d}y = 1, \tag{2.1}$$

$$\lim_{p\to\infty} \frac{\delta_p(x_1, y_1)}{\lim\limits_{(x,y)\to(0,0)} \delta_p(x, y)} = 0, \quad (x_1, y_1) \in \mathbb{R}^2 - \{(0, 0)\}. \tag{2.2}$$

The problem of this definition is what sense the limit is taken in. It is a limit of functions which is not a function.

We will show now an example of functions δ_p from Def. 2.20. Let us now introduce functions

$$\zeta_{(1),p}(t) = \begin{cases} \frac{p}{\pi} & \text{if } t = 0 \\ \frac{\sin pt}{\pi t} & \text{if } t \in \mathbb{R} - \{0\}, \end{cases} \qquad \zeta_{(2)}(s, t) = \begin{cases} \frac{s}{2\pi} & \text{if } t = 0, s \in \mathbb{R} \\ \frac{\sin st}{2\pi t} & \text{if } t \in \mathbb{R} - \{0\}, s \in \mathbb{R}, \end{cases} \tag{2.3}$$

Functions $\zeta_{(1),p}, \zeta_{(2)}$ are continuous since

$$\lim_{t\to 0} \zeta_{(1),p}(t) = \lim_{t\to 0} \frac{\sin pt}{\pi t} = \frac{p}{\pi}, \qquad \lim_{t\to 0} \zeta_{(2)}(s, t) = \lim_{t\to 0} \frac{\sin st}{2\pi t} = \frac{s}{2\pi}.$$

Furthermore

$$\zeta_{(2)}(p, t) - \zeta_{(2)}(-p, t) = \zeta_{(1),p}(t),$$

$$\frac{\mathrm{d}\zeta_{(2)}}{\mathrm{d}s}(s, t) = \frac{\cos st}{2\pi},$$

and $\zeta_{(1),p}(t) \in \mathcal{L}(\mathbb{R})$, especially

$$\int_{-\infty}^{\infty} \zeta_{(1),p}(t)\,\mathrm{d}t = 1.$$

For the proof of the last statement see Appendix A (page 92).

A graph of function $\zeta_{(1),p}$ is in Figure 2.1.

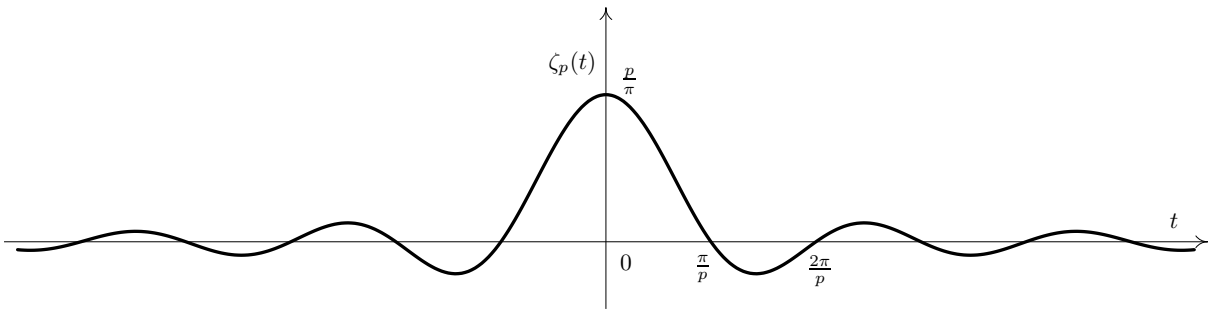


Figure 2.1: Graph of function $\zeta_p(t)$.

An example of δ_p functions is [34]

$$\begin{aligned}
\zeta_p(x,y) &= \frac{1}{4\pi^2}\int_{-p}^{p}\int_{-p}^{p} \mathrm{e}^{\mathrm{i}(sx+ty)}\,\mathrm{d}s\,\mathrm{d}t = \qquad (2.4)\\
&= \frac{1}{4\pi^2}\int_{-p}^{p}\int_{-p}^{p} (\cos(sx+ty)+\mathrm{i}\sin(sx+ty))\,\mathrm{d}s\,\mathrm{d}t =\\
&= \frac{1}{4\pi^2}\int_{-p}^{p}\int_{-p}^{p} (\cos sx\cos ty-\sin sx\sin ty+\mathrm{i}\sin sx\cos ty+\mathrm{i}\sin ty\cos sx))\,\mathrm{d}s\,\mathrm{d}t =\\
&= \frac{1}{4\pi^2}\int_{-p}^{p}\int_{-p}^{p} (\cos sx(\cos ty+\mathrm{i}\sin ty)+\sin sx(-\sin ty+\mathrm{i}\cos ty))\,\mathrm{d}s\,\mathrm{d}t =\\
&= \frac{1}{4\pi^2}\int_{-p}^{p}\cos sx\,\mathrm{d}s\int_{-p}^{p}(\cos ty+\mathrm{i}\sin ty)\,\mathrm{d}t +\\
&\quad +\frac{1}{4\pi^2}\int_{-p}^{p}\sin sx\,\mathrm{d}s\int_{-p}^{p}(-\sin ty+\mathrm{i}\cos ty)\,\mathrm{d}t.
\end{aligned}$$

Integral of a sine function over a symmetrical interval $(-p,p)$ is zero. Therefore

$$\zeta_p(x,y)=\frac{1}{4\pi^2}\int_{-p}^{p}\cos sx\,\mathrm{d}s\int_{-p}^{p}\cos ty\,\mathrm{d}t=\left[\zeta_{(2)}(s,x)\right]_{-p}^{p}\cdot\left[\zeta_{(2)}(t,y)\right]_{-p}^{p}=\zeta_{(1),p}(x)\zeta_{(1),p}(y).$$

Next we must prove that conditions (2.1), (2.2) hold for ζ_p.

Proof. Condition (2.1) holds not only for the limit $p\to\infty$, but also for each p as it is shown in [34].

$$\begin{aligned}
\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\zeta_p(x,y)\,\mathrm{d}x\,\mathrm{d}y &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\zeta_{(1),p}(x)\zeta_{(1),p}(y)\,\mathrm{d}x\,\mathrm{d}y = \int_{-\infty}^{\infty}\zeta_{(1),p}(x)\,\mathrm{d}x\int_{-\infty}^{\infty}\zeta_{(1),p}(y)\,\mathrm{d}y =\\
&= \frac{1}{\pi^2}2\int_{0}^{\infty}\frac{\sin px}{x}\,\mathrm{d}x\cdot 2\int_{0}^{\infty}\frac{\sin py}{y}\,\mathrm{d}y=\frac{4}{\pi^2}\frac{\pi}{2}\frac{\pi}{2}=1,
\end{aligned}$$

the pre-last equality is proved in Appendix A.

Condition (2.2) for ζ_p: Let $(x_1,y_1)\in\mathbb{R}^2-\{(0,0)\}$ arbitrarily. Then

$$\lim_{p\to\infty}\frac{\zeta_p(x_1,y_1)}{\lim\limits_{(x,y)\to(0,0)}\zeta_p(x,y)}=\lim_{p\to\infty}\frac{\zeta_p(x_1,y_1)}{\lim\limits_{(x,y)\to(0,0)}\zeta_{(1),p}(x)\zeta_{(1),p}(y)}=\lim_{p\to\infty}\frac{\zeta_p(x_1,y_1)}{\frac{p^2}{\pi^2}}=0.$$

□

2.3 The Fourier transform and the inverse Fourier transform

2.3.1 Functions in $\mathcal{L}(\mathbb{R})$

Definition 2.21 (Fourier transform of functions in $\mathcal{L}(\mathbb{R})$). Let $f(x) \in \mathcal{L}(\mathbb{R})$. The *Fourier transform* [15], [32] of function f is function $\mathcal{F}\{f\}(\xi) = F(\xi) : \mathbb{R} \to \mathbb{C}$ defined as

$$F(\xi) = \int_{-\infty}^{\infty} f(x)\mathrm{e}^{-\mathrm{i}x\xi}\,\mathrm{d}x.$$

Function F is also called the *Fourier spectrum* of function f.

Remark 2.22. Once the Fourier transform is defined for functions in $\mathcal{L}(\mathbb{R})$, it exists and is a bounded function as it is shown in [3]

$$|F(\xi)| = \left| \int_{-\infty}^{\infty} f(x)\mathrm{e}^{-\mathrm{i}x\xi}\,\mathrm{d}x \right| \leq \int_{-\infty}^{\infty} |f(x)|\,\mathrm{d}x < \infty.$$

Definition 2.23 (Inverse Fourier transform of functions in $\mathcal{L}(\mathbb{R})$). Let function $G(\xi) \in \mathcal{L}(\mathbb{R})$. The *inverse Fourier transform* [15], [32] of function G is function $\mathcal{F}^{-1}\{G\}(x) = g(x) : \mathbb{R} \to \mathbb{C}$ defined by

$$g(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} G(\xi)\mathrm{e}^{\mathrm{i}x\xi}\,\mathrm{d}\xi.$$

Remark 2.24. Similarly with the Fourier transform, the inverse Fourier transform exists and is a bounded function for every $G(\xi) \in \mathcal{L}(\mathbb{R})$. However, if we defined the inverse Fourier transform for functions which are Fourier spectra of functions in $\mathcal{L}(\mathbb{R})$, it may happen that $\mathcal{F}^{-1}\{\mathcal{F}\{f(x)\}\} \neq f(x)$ or the inverse Fourier transform is undefined. The most obvious example of $\mathcal{F}^{-1}\{\mathcal{F}\{f(x)\}\} \neq f(x)$ is taking a continuous function $f \in \mathcal{L}(\mathbb{R})$ and a function $g \in \mathcal{L}(\mathbb{R})$ which are equal almost everywhere (the notion "almost everywhere" to be defined in [9]). Then they have the same Fourier transform and if it exists and is equal to f, also the same inverse Fourier transform. The integration eliminates the discontinuity points in g and $\mathcal{F}^{-1}\{\mathcal{F}\{f(x)\}\} = f(x) \neq g(x)$. The case when the inverse Fourier transform is undefined is demonstrated by the following examples.

Example 2.25. Compute the Fourier transform of function

$$f(x) = \begin{cases} 1 & \text{if } -1 < x < 1 \\ 0 & \text{else.} \end{cases}$$

Solution. Obviously, $f \in \mathcal{L}(\mathbb{R})$.

$$F(\xi) = \int_{-1}^{1} \mathrm{e}^{-\mathrm{i}x\xi}\,\mathrm{d}x = \left[\frac{\mathrm{e}^{-\mathrm{i}x\xi}}{-\mathrm{i}\xi}\right]_{x=-1}^{1} = \frac{\mathrm{e}^{\mathrm{i}\xi} - \mathrm{e}^{-\mathrm{i}\xi}}{\mathrm{i}\xi} = 2\frac{\sin\xi}{\xi}.$$

The calculation is simplified in the sense that $\frac{\sin\xi}{\xi}$ is undefined for $\xi = 0$ and function $\zeta_{(1),1}(\xi)$ should be used instead. Then $F(\xi) = 2\pi\zeta_{(1),1}(\xi)$. □

Example 2.26. Compute the inverse Fourier transform of function $F(\xi) = 2\pi\zeta_{(1),1}(\xi)$ in $x = 1$, i.e. compute $\mathcal{F}^{-1}\left\{\mathcal{F}\left\{f(x)\right\}\right\}|_{x=1}$, where $f(x)$ is from Example 2.25.

Solution. We must check first if $F(\xi) \in \mathcal{L}(\mathbb{R})$. It is not, since

$$\int_{-\infty}^{\infty} \left|\frac{\sin x}{x}\right| \mathrm{d}x = 2\int_{0}^{\infty} \frac{|\sin x|}{x}\,\mathrm{d}x = \infty,$$

which is proved in Appendix B. Therefore, the inverse Fourier transform of $F(\xi)$ is undefined. □

Remark 2.27. We will show why the assumption that $F \in \mathcal{L}(\mathbb{R})$ cannot be omitted. We will try to compute the integral defining the inverse Fourier transform for $F(\xi) = 2\pi\zeta_{(1),1}(\xi)$, first for $x = 0$ and then for $x = 1$. For $x = 0$ we have

$$\frac{1}{2\pi}\int_{-\infty}^{\infty} 2\pi\zeta_{(1),1}(\xi)\mathrm{e}^{-\mathrm{i}0\xi}\,\mathrm{d}\xi = \frac{2}{\pi}\int_{0}^{\infty}\frac{\sin x}{x}\,\mathrm{d}x = 1 = f(0),$$

the pre-last equality is proved in Appendix A. For $x = 1$ we have

$$\begin{aligned}
&\frac{1}{2\pi}\int_{-\infty}^{\infty} 2\pi\zeta_{(1),1}(\xi)\mathrm{e}^{\mathrm{i}\xi\cdot 1}\,\mathrm{d}\xi = \int_{-\infty}^{0}\frac{\sin\xi}{\pi\xi}\mathrm{e}^{\mathrm{i}\xi}\,\mathrm{d}\xi + \int_{0}^{\infty}\frac{\sin\xi}{\pi\xi}\mathrm{e}^{\mathrm{i}\xi}\,\mathrm{d}\xi = \\
&= \int_{-\infty}^{0}\frac{\sin\xi}{\pi\xi}(\cos\xi + \mathrm{i}\sin\xi)\,\mathrm{d}\xi + \int_{0}^{\infty}\frac{\sin\xi}{\pi\xi}(\cos\xi + \mathrm{i}\sin\xi)\,\mathrm{d}\xi = \\
&= \int_{-\infty}^{0}\frac{\sin 2\xi}{2\pi\xi}\,\mathrm{d}\xi + \mathrm{i}\int_{-\infty}^{0}\frac{\sin^2\xi}{\pi\xi}\,\mathrm{d}\xi + \int_{0}^{\infty}\frac{\sin 2\xi}{2\pi\xi}\,\mathrm{d}\xi + \mathrm{i}\int_{0}^{\infty}\frac{\sin^2\xi}{\pi\xi}\,\mathrm{d}\xi
\end{aligned}$$

The first and the third term is equal to $\frac{1}{4}$ (for explanation see Appendix A on page 92). The fourth term is ∞ (for explanation see Appendix C on page 97) and in a similar way, the second one is $-\infty$. Therefore, the integral defining the inverse Fourier transform of $F(\xi) = 2\pi\zeta_{(1),1}(\xi)$ is undefined in $x = 1$.

This problem of integral convergence can be solved using the Cauchy principal value of the integral in the inverse Fourier transform.

Definition 2.28 (Cauchy principal value). Let f be a function $\mathbb{R} \to \mathbb{C}$. The value of

$$\lim_{r\to\infty}\int_{-r}^{r} f(x)\,\mathrm{d}x$$

is called the *Cauchy principal value* [7] of the integral

$$\int_{-\infty}^{\infty} f(x)\,\mathrm{d}x,$$

provided that this limit exists. We write

$$(\text{P.V.}) \int_{-\infty}^{\infty} f(x)\,\mathrm{d}x = \lim_{r\to\infty} \int_{-r}^{r} f(x)\,\mathrm{d}x.$$

Theorem 2.29 (Riemann-Lebesgue Lemma). If $f(x) \in \mathcal{L}(\mathbb{R})$, then [40]

$$\lim_{\xi\to\pm\infty} F(\xi) = 0.$$

Proof. The following proof is taken from [40] and adapted to complex functions.

Let us assume first that f is a rectangle function, i.e.

$$f(x) = \begin{cases} c & \text{if } a < x < b \\ 0 & \text{else} \end{cases}$$

for some $a, b, c \in \mathbb{R}, a < b$. Then

$$F(\xi) = \int_a^b c\mathrm{e}^{-\mathrm{i}x\xi}\,\mathrm{d}x = \left[\frac{c\mathrm{e}^{-\mathrm{i}x\xi}}{-\mathrm{i}\xi}\right]_{x=a}^{b}$$

which goes to zero as $\xi \to \pm\infty$. Now suppose that f is a step function, i.e. it is a linear combination of rectangle functions. Due to linearity of Riemann integration, the Fourier transform of f is a linear combination of Fourier transforms of these rectangle functions. By the definition of the Riemann integral, given f is real and $f \in \mathcal{L}(\mathbb{R})$, there exists a sequence f_n of step functions such that

$$\lim_{n\to\infty} \int_{-\infty}^{\infty} |f_n(x) - f(x)|\,\mathrm{d}x = 0.$$

Then for every $\xi \in \mathbb{R}$

$$\lim_{n\to\infty} |F_n(\xi) - F(\xi)| = \lim_{n\to\infty} \left| \int_{-\infty}^{\infty} f_n(x)\mathrm{e}^{-\mathrm{i}x\xi}\,\mathrm{d}x - \int_{-\infty}^{\infty} f(x)\mathrm{e}^{-\mathrm{i}x\xi}\,\mathrm{d}x \right|$$

and since both $f_n, f \in \mathcal{L}(\mathbb{R})$

$$\lim_{n\to\infty} |F_n(\xi) - F(\xi)| = \lim_{n\to\infty} \left| \int_{-\infty}^{\infty} (f_n(x) - f(x))\,\mathrm{e}^{-\mathrm{i}x\xi}\,\mathrm{d}x \right| \leq$$

$$\leq \lim_{n\to\infty} \left| \int_{-\infty}^{\infty} (f_n(x) - f(x))\,\mathrm{d}x \right| \leq \lim_{n\to\infty} \int_{-\infty}^{\infty} |f_n(x) - f(x)|\,\mathrm{d}x = 0$$

By the triangle inequality

$$F(\xi) \leq |F(\xi) - F_n(\xi)| + |F_n(\xi)|,$$

where the first term is bounded by any $\varepsilon > 0$ if n is big enough and the second term goes to zero if $\xi \to \pm\infty$. Therefore

$$\lim_{\xi\to\pm\infty} F(\xi) = 0.$$

If f is a complex function,

$$f(x) = u(x) + \mathrm{i}v(x),$$

where u and v are real functions of a real variable and $u, v \in \mathcal{L}(\mathbb{R})$. Assuming that $u(x), v(x)$ have Fourier spectra $U(\xi), V(\xi)$, we have

$$\lim_{\xi\to\pm\infty} U(\xi) = 0, \qquad \lim_{\xi\to\pm\infty} V(\xi) = 0,$$

therefore also

$$\lim_{\xi\to\pm\infty} F(\xi) = 0.$$

□

Definition 2.30 (C^1 functions). Function $f : \mathbb{R} \to \mathbb{C}$ is said to be of *class* C^1 [39] if it is continuous and both its real and imaginary part have continuous derivatives.

Definition 2.31 (Piecewise C^1 functions). Function $f(x) : \mathbb{R} \to \mathbb{C}$ is said to be *piecewise* C^1 [39] if it is continuous except for a finite number of discontinuity points in any bounded interval and both its real and imaginary part's derivatives are continuous except for a finite number of discontinuity points in any bounded interval. Furthermore, if $f(x)$, $(\mathfrak{Re} f(x))'$, or $(\mathfrak{Im} f(x))'$ is discontinuous in x_0, f, $(\mathfrak{Re} f(x))'$, and $(\mathfrak{Im} f(x))'$ have right- and left-hand limits in x_0.

Theorem 2.32 (Fourier Inversion Theorem for functions in $\mathcal{L}(\mathbb{R})$). If $f \in \mathcal{L}(\mathbb{R})$ and f is piecewise C^1, then [40]

$$\frac{1}{2\pi}(\mathrm{P.V.}) \int_{-\infty}^{\infty} F(\xi)\mathrm{e}^{\mathrm{i}x\xi}\,\mathrm{d}\xi = \lim_{r\to\infty} \frac{1}{2\pi} \int_{-r}^{r} F(\xi)\mathrm{e}^{\mathrm{i}x\xi}\,\mathrm{d}\xi = \frac{\lim_{t\to x+} f(x) + \lim_{t\to x-} f(x)}{2},$$

in particular, if f is continuous, then

$$\frac{1}{2\pi}(\mathrm{P.V.}) \int_{-\infty}^{\infty} F(\xi)\mathrm{e}^{\mathrm{i}x\xi}\,\mathrm{d}\xi = f(x).$$

Moreover, if also $F \in \mathcal{L}(\mathbb{R})$,

$$\mathcal{F}^{-1}\left\{\mathcal{F}\left\{f(x)\right\}\right\} = \frac{\lim_{t\to x+} f(x) + \lim_{t\to x-} f(x)}{2},$$

and for a continuous function f with $F \in \mathcal{L}(\mathbb{R})$ we have

$$\mathcal{F}^{-1}\left\{\mathcal{F}\left\{f(x)\right\}\right\} = f(x).$$

Proof. The proof is taken from [40] and revised for complex functions and the consequence for functions in $\mathcal{L}(\mathbb{R})$ is added.

$$\frac{1}{2\pi}\int_{-r}^{r} F(\xi)\mathrm{e}^{\mathrm{i}x\xi}\,\mathrm{d}\xi = \frac{1}{2\pi}\int_{-r}^{r}\left(\int_{-\infty}^{\infty} f(s)\mathrm{e}^{-\mathrm{i}s\xi}\,\mathrm{d}s\right)\mathrm{e}^{\mathrm{i}x\xi}\,\mathrm{d}\xi$$

Since $f \in \mathcal{L}(\mathbb{R})$,

$$\frac{1}{2\pi}\int_{-r}^{r}\left(\int_{-\infty}^{\infty} f(s)\mathrm{e}^{-\mathrm{i}s\xi}\,\mathrm{d}s\right)\mathrm{e}^{\mathrm{i}x\xi}\,\mathrm{d}\xi = \frac{1}{2\pi}\int_{-\infty}^{\infty} f(s)\left(\int_{-r}^{r}\mathrm{e}^{\mathrm{i}(x-s)\xi}\,\mathrm{d}\xi\right)\,\mathrm{d}s =$$
$$= \frac{1}{\pi}\int_{-\infty}^{\infty} f(s)\frac{\sin(r(x-s))}{x-s}\,\mathrm{d}s = \frac{1}{\pi}\int_{-\infty}^{\infty} f(x-s)\frac{\sin rs}{s}\,\mathrm{d}s.$$

Let us introduce shorter notation for right- and left-hand limits:

$$f(x^-) = \lim_{t\to x-} f(t), \qquad f(x^+) = \lim_{t\to x+} f(t).$$

Then we can rewrite $f(x-s)$ as

$$f(x-s) = \begin{cases} f(x^+) + (f(x-s) - f(x^+)) & \text{if } s < 0 \\ f(x^-) + (f(x-s) - f(x^-)) & \text{if } s \geq 0. \end{cases}$$

This is necessary for treating the case if f is not continuous in x. As the value in a single point does not matter for a Riemann integral, it does not make any difference how the case $s = 0$ is treated. It can be added to any of the cases above. Then the last integral can be written as

$$\begin{aligned}\frac{1}{\pi}\int_{-\infty}^{\infty} f(x-s)\frac{\sin rs}{s}\,\mathrm{d}s \;&=\; \frac{1}{\pi}\int_{-\infty}^{0}\left\{f(x^+) + (f(x-s) - f(x^+))\right\}\frac{\sin rs}{s}\,\mathrm{d}s + \\ &\quad \frac{1}{\pi}\int_{0}^{\infty}\left\{f(x^-) + (f(x-s) - f(x^-))\right\}\frac{\sin rs}{s}\,\mathrm{d}s.\end{aligned}$$

$$\frac{1}{\pi}\int_{0}^{\infty} f(x^-)\frac{\sin rs}{s}\,\mathrm{d}s = \frac{f(x^-)}{\pi}\int_{0}^{\infty}\frac{\sin rs}{s}\,\mathrm{d}s = \frac{f(x^-)}{\pi}\cdot\frac{\pi}{2} = \frac{f(x^-)}{2}$$

For the pre-last equality see Appendix A (page 92). Similarly,

$$\frac{1}{\pi}\int_{0}^{\infty} f(x^+)\frac{\sin rs}{s}\,\mathrm{d}s = \frac{f(x^+)}{2}.$$

It remains to prove that the rest of the terms goes to zero as r goes to infinity.

$$\frac{1}{\pi}\int_0^\infty (f(x-s)-f(x^-))\frac{\sin rs}{s}\,\mathrm{d}s =$$

$$= \underbrace{\frac{1}{\pi}\int_0^K (f(x-s)-f(x^-))\frac{\sin rs}{s}\,\mathrm{d}s}_{\text{denote by } a(x)} + \underbrace{\frac{1}{\pi}\int_K^\infty (f(x-s)-f(x^-))\frac{\sin rs}{s}\,\mathrm{d}s}_{\text{denote by } b(x)}$$

for fixed $K > 0$. Now set

$$g(s) = \chi_{\langle 0,K)}(s)\frac{f(x-s)-f(x^-)}{s}$$

for fixed x. Then

$$a(x) = \frac{1}{\pi}\int_{-\infty}^\infty g(s)\sin rs\,\mathrm{d}s = \frac{1}{\pi}\int_{-\infty}^\infty g(s)\frac{\mathrm{e}^{\mathrm{i}rs}-\mathrm{e}^{-\mathrm{i}rs}}{2\mathrm{i}}\,\mathrm{d}s = \frac{1}{2\pi\mathrm{i}}(G(-r)-G(r)),$$

where G is the Fourier transform of g. Since $g \in \mathcal{L}(\mathbb{R})$, by the Riemann-Lebesgue Lemma 2.29 (page 14) we have $a(x) \to 0$ as $r \to \infty$. Since $|\sin rs| \le 1$,

$$\begin{aligned} |b(x)| &\le \frac{1}{\pi}\left|\int_K^\infty f(x-s)\frac{1}{s}\,\mathrm{d}s\right| + \frac{1}{\pi}\left|\int_K^\infty f(x^-)\frac{\sin rs}{s}\,\mathrm{d}s\right| = \left|\begin{array}{rcl} rs &=& t \\ r\,\mathrm{d}s &=& \mathrm{d}t \end{array}\right| \le \\ &\le \frac{1}{\pi}\int_K^\infty |f(x-s)|\frac{1}{K}\,\mathrm{d}s + \frac{|f(x^-)|}{\pi}\left|\int_{Kr}^\infty \frac{\sin t}{t}\,\mathrm{d}t\right| \end{aligned}$$

Since $f \in \mathcal{L}(\mathbb{R})$,

$$\int_K^\infty |f(x-s)|\,\mathrm{d}s \to 0 \text{ as } K \to \infty,$$

therefore the first term of $|b(x)|$ goes to zero as K goes to infinity. The fact that

$$\int_0^\infty \frac{\sin x}{x}\,\mathrm{d}x = \frac{\pi}{2} < \infty$$

implies that

$$\int_{Kr}^\infty \frac{\sin t}{t}\,\mathrm{d}t \to 0 \text{ as } K \to \infty.$$

Thus, if we choose any $\varepsilon > 0$, we can find K big enough such that $|b(x)| < \varepsilon$ for all $x \in \mathbb{R}$. We have proved that

$$\frac{1}{\pi}\int_0^\infty (f(x-s)-f(x^-))\frac{\sin rs}{s}\,\mathrm{d}s$$

goes to zero as r goes to infinity. In a similar way, we can prove that

$$\frac{1}{\pi}\int_0^\infty (f(x-s)-f(x^+))\frac{\sin rs}{s}\,\mathrm{d}s$$

goes to zero as r goes to infinity. We can conclude that

$$\lim_{r\to\infty}\frac{1}{2\pi}\int_{-r}^{r} F(\xi)\mathrm{e}^{\mathrm{i}x\xi}\,\mathrm{d}\xi = \frac{\lim_{t\to x+} f(x)+\lim_{t\to x-} f(x)}{2}.$$

If the Fourier spectrum $F \in \mathcal{L}(\mathbb{R})$, we can use the reasoning from Remark 2.22 on page 12, this time for the inverse Fourier transform. We obtain that for every $x \in \mathbb{R}$, $|\mathcal{F}^{-1}\{F\}(x)|$ exists and is a finite number. Therefore,

$$\lim_{r\to\infty}\frac{1}{2\pi}\int_{-r}^{r} F(\xi)\mathrm{e}^{\mathrm{i}x\xi}\,\mathrm{d}\xi = \frac{1}{2\pi}\int_{-\infty}^{\infty} F(\xi)\mathrm{e}^{\mathrm{i}x\xi}\,\mathrm{d}\xi = \mathcal{F}^{-1}\{F(\xi)\} = \frac{\lim_{t\to x+} f(x)+\lim_{t\to x-} f(x)}{2}.$$

Finally, if f is continuous and $F \in \mathcal{L}(\mathbb{R})$,

$$\mathcal{F}^{-1}\{\mathcal{F}\{f(x)\}\} = f(x).$$

□

Example 2.33. Compute

$$\frac{1}{2\pi}(\text{P.V.})\int_{-\infty}^{\infty} F(\xi)\mathrm{e}^{\mathrm{i}x\xi}\,\mathrm{d}\xi$$

for function $F(\xi) = 2\pi\zeta_{(1),1}(\xi)$ from Example 2.25 on page 12.

Solution. The Fourier Inversion Theorem (Thm. 2.32) says that

$$\frac{1}{2\pi}(\text{P.V.})\int_{-\infty}^{\infty} F(\xi)\mathrm{e}^{\mathrm{i}x\xi}\,\mathrm{d}\xi = \frac{1}{2\pi}(\text{P.V.})\int_{-\infty}^{\infty} 2\pi\zeta_{(1),1}(\xi)\mathrm{e}^{\mathrm{i}x\xi}\,\mathrm{d}\xi = \begin{cases} 1 & \text{if } -1<x<1 \\ \frac{1}{2} & \text{if } x\in\{-1,1\} \\ 0 & \text{else.} \end{cases}$$

□

2.3.2 Functions in $\mathcal{L}(\mathbb{R}^2)$

Definition 2.34 (Fourier transform of functions in $\mathcal{L}(\mathbb{R}^2)$). Let $f(x,y) \in \mathcal{L}(\mathbb{R}^2)$. The *Fourier transform* [8] of function f is function $\mathcal{F}\{f\}(\xi,\eta) = F(\xi,\eta) : \mathbb{R}^2 \to \mathbb{C}$ defined as

$$F(\xi,\eta) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\mathrm{e}^{-\mathrm{i}(x\xi+y\eta)}\,\mathrm{d}x\,\mathrm{d}y.$$

Function F is also called the *Fourier spectrum* of function f.

Definition 2.35 (Inverse Fourier transform of functions in $\mathcal{L}(\mathbb{R}^2)$). Let function $G(\xi,\eta) \in \mathcal{L}(\mathbb{R}^2)$. The *inverse Fourier transform* [8] of function G is function $\mathcal{F}^{-1}\{G\}(x,y) = g(x,y) : \mathbb{R}^2 \to \mathbb{C}$ defined as

$$g(x,y) = \frac{1}{4\pi^2} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} G(\xi,\eta) \mathrm{e}^{\mathrm{i}(x\xi+y\eta)} \,\mathrm{d}\xi\,\mathrm{d}\eta.$$

Remark 2.36. The use of δ_p functions for definition of the δ-distribution (Definition 2.20 on page 10) used e.g. in [34] is quite unusual in literature and leads to the following claim called the Fundamental Theorem on the Fourier Transform in [34]. We sketch what is wrong with it and try to suggest a solution.

Let function $f(x,y) \in \mathcal{L}(\mathbb{R}^2)$. Then for every $(x,y) \in \mathbb{R}^2$ where f is continuous it holds

$$\mathcal{F}^{-1}\{\mathcal{F}\{f(x,y)\}\} = f(x,y).$$

In discontinuity points, $\mathcal{F}^{-1}\{\mathcal{F}\{f(x,y)\}\}$ equals to the average of f over an infinitesimal neighborhood of the discontinuity point.

The following can be found in [34] as a proof of the claim:

$$\begin{aligned}
\mathcal{F}^{-1}\{\mathcal{F}\{f(x,y)\}\} &= \frac{1}{4\pi^2} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(s,t) \mathrm{e}^{-\mathrm{i}(\xi s+\eta t)} \,\mathrm{d}s\,\mathrm{d}t \right) \mathrm{e}^{\mathrm{i}(\xi x+\eta y)} \,\mathrm{d}\xi\,\mathrm{d}\eta = \\
&= \frac{1}{4\pi^2} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(s,t) \left(\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \mathrm{e}^{\mathrm{i}[\xi(x-s)+\eta(y-t)]} \,\mathrm{d}\xi\,\mathrm{d}\eta \right) \mathrm{d}s\,\mathrm{d}t.
\end{aligned}$$

The inner integral is function ζ_p from formula (2.4) on page 11 for $p \to \infty$ shifted by (s,t) and multiplied by $4\pi^2$, therefore we can write

$$\begin{aligned}
\mathcal{F}^{-1}\{\mathcal{F}\{f(x,y)\}\} &= \frac{1}{4\pi^2} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(s,t) 4\pi^2 \delta(x-s, y-t) \,\mathrm{d}s\,\mathrm{d}t = \\
&= \frac{1}{4\pi^2} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(s,t) 4\pi^2 \delta(s-x, t-y) \,\mathrm{d}s\,\mathrm{d}t,
\end{aligned}$$

which equals the average of f over an infinitesimal neighborhood of the discontinuity point.

The biggest drawback of the statement is that it allows to compute the inverse Fourier transform of Fourier spectra $F(\xi,\eta)$ which are not in $\mathcal{L}(\mathbb{R}^2)$. This may not be a problem of this statement, but rather on the concept of improper double integral itself. It is not unified in literature whether an improper double integral is defined in the way I define it in Definition 2.7 on page 7 or as the principal value of that integral (used e.g. in [38]). Unfortunately, Komrska in [34] does not say which definition he is using.

Another drawback of the statement is the "infinitesimal neighborhood" and its shape. The average value of f over an infinitesimal neighborhood of a discontinuity point may depend on the shape of the neighborhood. Consider the case illustrated in Fig. 2.2.

Function $f(x, y) = 1$ on the gray area (continuing upwards) and is equal to zero otherwise. Let us focus on the neighborhoods of the discontinuity point B and consider two types of neighborhoods, an ellipse and a square. Independently on the size of the square, the average value of f over the square neighborhood of B is $\frac{1}{4}$ since f equals to 1 on a fourth of the square. The size of the area of the ellipse where $f(x, y) = 1$ depends on the shape of the ellipse, namely on its elongation. In the case depicted in Fig. 2.2, the area is less than $\frac{1}{4}$. From the definition of function ζ_p, the neighborhood is square-shaped.

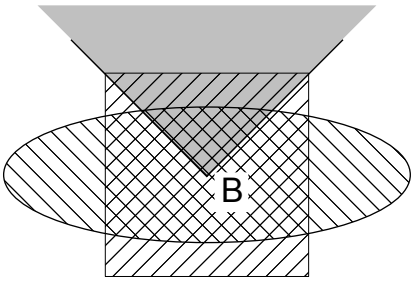


Figure 2.2: Illustration of the fact that the average value of function f over a neighborhood of a discontinuity point B may depend on the shape of the neighborhood.

In one dimension, the formulation of the Fourier Inversion Theorem was quite straightforward using the Cauchy principal value. In more dimensions, it is more complicated. One approach is to introduce a cut-off function which simulates the effect of the Cauchy principal value used in one dimension.

Theorem 2.37 (Fourier Inversion Theorem for functions in $\mathcal{L}(\mathbb{R}^2)$). If $f(\xi, \eta) \in \mathcal{L}(\mathbb{R}^2)$ and is continuous on $\mathbb{R}^2$, then for every $(\xi, \eta) \in \mathbb{R}^2$ [26]

$$f(x, y) = \lim_{\epsilon \to 0} \frac{1}{4\pi^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} F(\xi, \eta) \mathrm{e}^{\mathrm{i}(x\xi + y\eta)} \mathrm{e}^{-\epsilon^2 \frac{\xi^2 + \eta^2}{2}} \, \mathrm{d}\xi \, \mathrm{d}\eta.$$

If also $F(\xi, \eta) \in \mathcal{L}(\mathbb{R}^2)$, then

$$\mathcal{F}^{-1}\{\mathcal{F}\{f(x, y)\}\} = \frac{1}{4\pi^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} F(\xi, \eta) \mathrm{e}^{\mathrm{i}(x\xi + y\eta)} \, \mathrm{d}\xi \, \mathrm{d}\eta = f(x, y).$$

Proof. A proof together with a more general derivation can be found in [46]. □

Theorem 2.38. Let $f \in \mathcal{L}(\mathbb{R}^2)$ such that $f(x, y) = g(x)h(y)$, $g, h \in \mathcal{L}(\mathbb{R})$ and g and h are piecewise C^1. Let $F(\xi, \eta)$ be the Fourier transform of f. If the following limit exists

$$\lim_{(s,t) \to (x,y)} f(s, t),$$

then

$$\lim_{r \to \infty} \frac{1}{4\pi^2} \int_{-r}^{r} \int_{-r}^{r} F(\xi, \eta) \mathrm{e}^{\mathrm{i}(x\xi + y\eta)} \, \mathrm{d}\xi \, \mathrm{d}\eta =$$

$$\frac{1}{4\pi^2} (\mathrm{P.V.}) \int_{-\infty}^{\infty} \left((\mathrm{P.V.}) \int_{-\infty}^{\infty} F(\xi, \eta) \mathrm{e}^{\mathrm{i}(x\xi + y\eta)} \, \mathrm{d}\xi \right) \mathrm{d}\eta = \lim_{(s,t) \to (x,y)} f(s, t).$$

Proof.

$$\lim_{r\to\infty}\frac{1}{4\pi^2}\int_{-r}^{r}\int_{-r}^{r}F(\xi,\eta)\mathrm{e}^{\mathrm{i}(x\xi+y\eta)}\,\mathrm{d}\xi\,\mathrm{d}\eta =$$

$$= \lim_{r\to\infty}\frac{1}{4\pi^2}\int_{-r}^{r}\int_{-r}^{r}\left(\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}f(s,t)\mathrm{e}^{-\mathrm{i}(s\xi+t\eta)}\,\mathrm{d}s\,\mathrm{d}t\right)\mathrm{e}^{\mathrm{i}(x\xi+y\eta)}\,\mathrm{d}\xi\,\mathrm{d}\eta =$$

$$= \lim_{r\to\infty}\frac{1}{4\pi^2}\int_{-r}^{r}\int_{-r}^{r}\left(\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}g(s)h(t)\mathrm{e}^{-\mathrm{i}(s\xi+t\eta)}\,\mathrm{d}s\,\mathrm{d}t\right)\mathrm{e}^{\mathrm{i}(x\xi+y\eta)}\,\mathrm{d}\xi\,\mathrm{d}\eta =$$

$$= \lim_{r\to\infty}\frac{1}{2\pi}\int_{-r}^{r}\left(\int_{-\infty}^{\infty}g(s)\mathrm{e}^{\mathrm{i}\xi(x-s)}\,\mathrm{d}s\right)\mathrm{d}\xi\cdot\frac{1}{2\pi}\int_{-r}^{r}\left(\int_{-\infty}^{\infty}h(t)\mathrm{e}^{\mathrm{i}\eta(y-t)}\,\mathrm{d}t\right)\mathrm{d}\eta$$

By the Fourier Inversion Theorem for functions in $\mathcal{L}(\mathbb{R})$ (Thm. 2.32, page 15), we have

$$\lim_{r\to\infty}\left(\frac{1}{2\pi}\int_{-r}^{r}\int_{-\infty}^{\infty}g(s)\mathrm{e}^{\mathrm{i}\xi(x-s)}\,\mathrm{d}\xi\,\mathrm{d}s\frac{1}{2\pi}\int_{-r}^{r}\int_{-\infty}^{\infty}h(t)\mathrm{e}^{\mathrm{i}\eta(y-t)}\,\mathrm{d}\eta\,\mathrm{d}t\right) =$$
$$= \lim_{s\to x}g(s)\lim_{t\to y}h(t) = \lim_{(s,t)\to(x,y)}f(s,t).$$

□

Definition 2.39 (Amplitude spectrum, phase spectrum). Let function $f(x,y)\in\mathcal{L}(\mathbb{R}^2)$ have Fourier spectra $F(\xi,\eta)$. The *amplitude spectrum* [8] of function f is a function $A(\xi,\eta):\mathbb{R}^2\to\mathbb{R}_0^+$ defined as

$$A(\xi,\eta) = |\mathcal{F}\{f(x,y)\}| = |F(\xi,\eta)|.$$

The *phase spectrum* [8] of function f is a function $\Phi(\xi,\eta):\mathbb{R}^2\to\langle 0,2\pi)$ defined as

$$\begin{aligned}\mathfrak{Re}F(\xi,\eta) &= A(\xi,\eta)\cos\Phi(\xi,\eta),\\ \mathfrak{Im}F(\xi,\eta) &= A(\xi,\eta)\sin\Phi(\xi,\eta).\end{aligned}$$

If $A(\xi,\eta)=0$ for some (ξ,η), we define $\Phi(\xi,\eta)=0$.

2.4 Properties of the Fourier transform

Table 2.1 shows the basic properties of the Fourier transform. It is easy to prove them or they are proved below. Functions are listed on the left, their Fourier transforms on the right. $f(x,y), g(x,y)$ are functions in $\mathcal{L}(\mathbb{R}^2)$, $\alpha,\beta,x_0,y_0,\xi_0,\eta_0$ are real constants, $f*g$ denotes convolution of functions f,g (defined in Section 2.5 on page 31). The formulae are taken from [8] and [26] (adapted to two dimensions).

The second formula is linearity of the Fourier transform, the third formula is the Shift Theorem (Thm. 2.42), the fourth one is called the Modulation Theorem [8], the fifth one is the Scale-change Theorem (Thm. 2.43), the next one is a consequence of the Scale-change Theorem (Corollary 2.44). The last three formulae are proved in Theorems 2.53, 2.60 and 2.61.

1.	$f(x,y)$	$F(\xi,\eta)$
2.	$\alpha f(x,y)+\beta g(x,y)$	$\alpha F(\xi,\eta)+\beta G(\xi,\eta)$
3.	$f(x-x_0,y-y_0)$	$F(\xi,\eta)\mathrm{e}^{-\mathrm{i}(\xi x_0+\eta y_0)}$
4.	$\mathrm{e}^{\mathrm{i}(\xi_0 x+\eta_0 y)}f(x,y)$	$F(\xi-\xi_0,\eta-\eta_0)$
5.	$f(\alpha x,\alpha y)$	$\frac{1}{\alpha^2}F\left(\frac{\xi}{\alpha},\frac{\eta}{\alpha}\right)$
6.	$f(-x,-y)$	$F(-\xi,-\eta)$
7.	$f^*(x,y)$	$F^*(-\xi,-\eta)$
8.	$(f*g)(x,y)$	$F(\xi,\eta)G(\xi,\eta)$
9.	$f(x,y)g(x,y)$	$\frac{1}{4\pi^2}(F*G)(\xi,\eta)$

Table 2.1: Basic properties of the Fourier transform.

Definition 2.40 (Integral part). Let $a\in\mathbb{R}$. The *integral part* $\lfloor a\rfloor$ of the real number a is defined as [24]

$$\lfloor a\rfloor=\max\{n\in\mathbb{Z}:n\le a\}.$$

Definition 2.41 (Addition modulo 2π). Let $a,b\in\mathbb{R}$. The binary operation $\mathbb{R}^2\to\langle 0,2\pi)$ denoted $\oplus$ is defined as

$$a\oplus b\equiv a+b\ (\mathrm{mod}\ 2\pi)=(a+b)-2\pi\left\lfloor\frac{a+b}{2\pi}\right\rfloor$$

and is called *addition modulo* 2π [24].

Theorem 2.42 (Shift Theorem). Let $f_1(x,y)\in\mathcal{L}(\mathbb{R}^2)$ and let $F_1(\xi,\eta)$ be its Fourier spectrum. Let us suppose function

$$f_2(x,y)=f_1(x-x_0,y-y_0),$$

where $x_0,y_0\in\mathbb{R}$ are given constants. Let $F_2(\xi,\eta)$ be the Fourier spectrum of function $f_2(x,y)$. Then it holds (partially from [8])

$$F_2(\xi,\eta)=F_1(\xi,\eta)\mathrm{e}^{-\mathrm{i}(\xi x_0+\eta y_0)},$$

$$A_2(\xi,\eta)=A_1(\xi,\eta),$$

$$\Phi_2(\xi,\eta)=\Phi_1(\xi,\eta)\oplus(-\xi x_0-\eta y_0).$$

Proof. Proof taken from [3] and modified for two dimensions.

$$F_2(\xi,\eta)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}f_1(x-x_0,y-y_0)\mathrm{e}^{-\mathrm{i}(x\xi+y\eta)}\,\mathrm{d}x\,\mathrm{d}y=$$

$$=\left|\begin{array}{llll}s=x-x_0 & x=s+x_0 & \mathrm{d}s=\mathrm{d}x\\ t=y-y_0 & y=t+y_0 & \mathrm{d}t=\mathrm{d}y\end{array}\right|=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}f_1(s,t)\mathrm{e}^{-\mathrm{i}(\xi(s+x_0)+\eta(t+y_0))}\,\mathrm{d}s\,\mathrm{d}t=$$

$$=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}f_1(s,t)\mathrm{e}^{-\mathrm{i}(\xi s+\eta t)}\mathrm{e}^{-\mathrm{i}(\xi x_0+\eta y_0)}\,\mathrm{d}s\,\mathrm{d}t=F_1(\xi,\eta)\mathrm{e}^{-\mathrm{i}(\xi x_0+\eta y_0)}$$

Let A_1, A_2 be the amplitude spectra of functions f_1, f_2. Then

$$A_2(\xi,\eta) = \left|F_1(\xi,\eta)\mathrm{e}^{-\mathrm{i}(\xi x_0+\eta y_0)}\right| = |F_1(\xi,\eta)| \cdot \left|\mathrm{e}^{-\mathrm{i}(\xi x_0+\eta y_0)}\right| = A_1(\xi,\eta)\cdot 1 = A_1(\xi,\eta).$$

For the purpose of the phase spectrum, spectrum F_2 can be written as

$$F_2(\xi,\eta) = F_1(\xi,\eta)\mathrm{e}^{-\mathrm{i}(\xi x_0+\eta y_0)} = F_1(\xi,\eta)(\cos(\xi x_0+\eta y_0) + \mathrm{i}\sin(\xi x_0+\eta y_0))$$

and Moivre's theorem [18] implies that

$$\Phi_2(\xi,\eta) = \Phi_1(\xi,\eta) \oplus (-\xi x_0 - \eta y_0).$$

□

Theorem 2.43 (Scale-change Theorem). Let $f_1(x,y) \in \mathcal{L}(\mathbb{R}^2)$ and let $F_1(\xi,\eta)$ be its Fourier spectrum. Let us suppose function

$$f_2(x,y) = f_1(\alpha x, \alpha y),$$

where $\alpha \in \mathbb{R}^+$ is a given constant. Let $F_2(\xi,\eta)$ be the Fourier spectrum of function $f_2(x,y)$. Let $A_1(\xi,\eta), A_2(\xi,\eta)$ be the amplitude spectra and $\Phi_1(\xi,\eta), \Phi_2(\xi,\eta)$ the phase spectra of functions f_1, f_2. Then it holds (partially from [8])

$$F_2(\xi,\eta) = \frac{1}{\alpha^2}F_1\left(\frac{\xi}{\alpha},\frac{\eta}{\alpha}\right),$$
$$A_2(\xi,\eta) = \frac{1}{\alpha^2}A_1\left(\frac{\xi}{\alpha},\frac{\eta}{\alpha}\right),$$
$$\Phi_2(\xi,\eta) = \Phi_1\left(\frac{\xi}{\alpha},\frac{\eta}{\alpha}\right).$$

Proof. Proof taken from [3] and modified for two dimensions.

$$\begin{aligned} F_2(\xi,\eta) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_2(x,y)\mathrm{e}^{-\mathrm{i}(x\xi+y\eta)}\,\mathrm{d}x\,\mathrm{d}y = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_1(\alpha x,\alpha y)\mathrm{e}^{-\mathrm{i}(x\xi+y\eta)}\,\mathrm{d}x\,\mathrm{d}y = \\ &= \left|\begin{array}{ccccccccc} s &=& \alpha x & x &=& \frac{s}{\alpha} & \mathrm{d}x &=& \frac{\mathrm{d}s}{\alpha} \\ t &=& \alpha y & y &=& \frac{t}{\alpha} & \mathrm{d}y &=& \frac{\mathrm{d}t}{\alpha} \end{array}\right| = \\ &= \frac{1}{\alpha^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_2(s,t)\mathrm{e}^{-\mathrm{i}\left(\frac{s}{\alpha}\xi+\frac{t}{\alpha}\eta\right)}\,\mathrm{d}s\,\mathrm{d}t = \frac{1}{\alpha^2}F_1\left(\frac{\xi}{\alpha},\frac{\eta}{\alpha}\right) \end{aligned}$$

$$A_2(\xi,\eta) = |F_2(\xi,\eta)| = \frac{1}{\alpha^2}\left|F_1\left(\frac{\xi}{\alpha},\frac{\eta}{\alpha}\right)\right| = \frac{1}{\alpha^2}A_1\left(\frac{\xi}{\alpha},\frac{\eta}{\alpha}\right)$$

The Moivre's theorem [18] implies that

$$\Phi_2(\xi,\eta) = \Phi_1\left(\frac{\xi}{\alpha},\frac{\eta}{\alpha}\right).$$

□

Corollary 2.44. Let $f(x,y) \in \mathcal{L}(\mathbb{R}^2)$ and let $F(\xi,\eta)$ be its Fourier spectrum. Then [8]

$$\mathcal{F}\{f(-x,-y)\} = F(-x,-y).$$

Proof. By the Scale-change Theorem for $\alpha = -1$. □

Theorem 2.45 (Rotation Theorem). Let $f_1(x,y) \in \mathcal{L}(\mathbb{R}^2)$ and let $F_1(\xi,\eta)$ be its Fourier spectrum. Let us suppose function

$$f_2(x,y) = f_1(x\cos\theta - y\sin\theta, x\sin\theta + y\cos\theta), \tag{2.5}$$

where $\theta \in \langle 0, 2\pi)$ is a given constant (see Fig. 2.3). Let $F_2(\xi,\eta)$ be the Fourier spectrum of function $f_2(x,y)$. Let $A_1(\xi,\eta), A_2(\xi,\eta)$ be the amplitude spectra and $\Phi_1(\xi,\eta), \Phi_2(\xi,\eta)$ the phase spectra of functions f_1, f_2. Then it holds (partially taken from [45]):

$$\begin{aligned}
F_2(\xi,\eta) &= F_1(\xi\cos\theta - \eta\sin\theta, \xi\sin\theta + \eta\cos\theta),\\
A_2(\xi,\eta) &= A_1(\xi\cos\theta - \eta\sin\theta, \xi\sin\theta + \eta\cos\theta),\\
\Phi_2(\xi,\eta) &= \Phi_1(\xi\cos\theta - \eta\sin\theta, \xi\sin\theta + \eta\cos\theta).
\end{aligned}$$

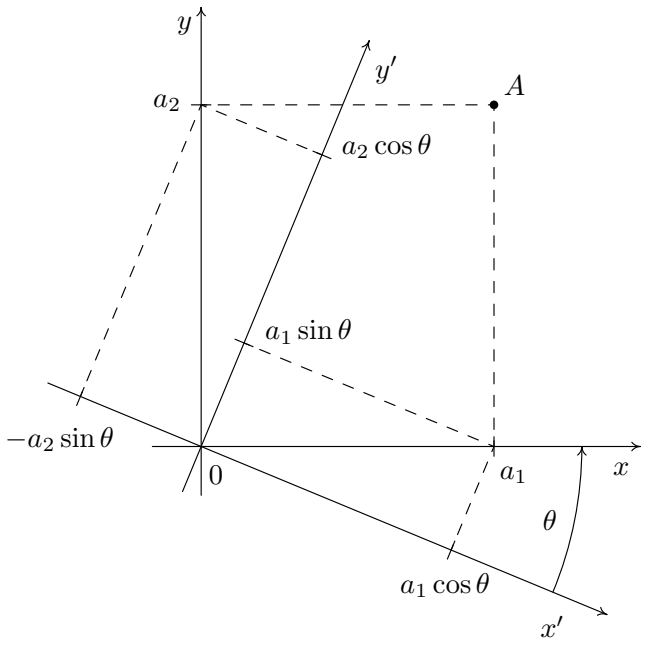


Figure 2.3: Derivation of the rotation formula (2.5). Point A has coordinates (a_1, a_2) in coordinate system x, y, which is rotated with respect to coordinate system x', y' by angle θ. Point $(a_1, 0)$ in x, y has coordinates $(a_1\cos\theta, a_1\sin\theta)$ in x', y'. Similarly, point $(0, a_2)$ in x, y has coordinates $(-a_2\sin\theta, a_2\cos\theta)$ in x', y'. Thus, point $A = (a_1, a_2)$ in coordinate system x, y has coordinates $(a_1\cos\theta - a_2\sin\theta, a_1\sin\theta + a_2\cos\theta)$ in coordinate system x', y'.

Proof.

$$\begin{aligned}
F_2(\xi,\eta) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_2(x,y)\mathrm{e}^{-\mathrm{i}(x\xi+y\eta)}\,\mathrm{d}x\,\mathrm{d}y = \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_1(x\cos\theta - y\sin\theta, x\sin\theta + y\cos\theta)\mathrm{e}^{-\mathrm{i}(x\xi+y\eta)}\,\mathrm{d}x\,\mathrm{d}y = \\
&= \left|\begin{array}{llll} s &= x\cos\theta - y\sin\theta & x &= s\cos\theta + t\sin\theta \\ t &= x\sin\theta + y\cos\theta & y &= -s\sin\theta + t\cos\theta \end{array}\right| = \\
&= \left|\begin{array}{cc} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{array}\right| \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_1(s,t)\mathrm{e}^{-\mathrm{i}[\xi(s\cos\theta+t\sin\theta)+\eta(-s\sin\theta+t\cos\theta)]}\,\mathrm{d}s\,\mathrm{d}t = \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_1(s,t)\mathrm{e}^{-\mathrm{i}[s(\xi\cos\theta-\eta\sin\theta)+t(\xi\sin\theta+\eta\cos\theta)]}\,\mathrm{d}s\,\mathrm{d}t = \\
&= F_1(\xi\cos\theta - \eta\sin\theta, \xi\sin\theta + \eta\cos\theta)
\end{aligned}$$

Formulae for the amplitude and phase spectra are clear from the formula for the Fourier spectrum. □

Theorem 2.46 (Similarity Theorem). Let $f_1(x,y) \in \mathcal{L}(\mathbb{R}^2)$ and let $F_1(\xi,\eta)$ be its Fourier spectrum. Let us suppose function

$$f_2(x,y) = f_1(\alpha x\cos\theta - \alpha y\sin\theta - x_0, \alpha x\sin\theta + \alpha y\cos\theta - y_0),$$

where $\theta \in \langle 0, 2\pi)$, $\alpha \in \mathbb{R}^+$, $x_0, y_0 \in \mathbb{R}$ are given constants. Let $F_2(\xi,\eta)$ be the Fourier spectrum of function $f_2(x,y)$. Let $A_1(\xi,\eta), A_2(\xi,\eta)$ be the amplitude spectra and $\Phi_1(\xi,\eta)$, $\Phi_2(\xi,\eta)$ the phase spectra of functions f_1, f_2. Then it holds:

$$\begin{aligned}
F_2(\xi,\eta) &= \frac{1}{\alpha^2}\mathrm{e}^{-\mathrm{i}(\xi x_0+\eta y_0)} F_1\left(\frac{\xi}{\alpha}\cos\theta - \frac{\eta}{\alpha}\sin\theta, \frac{\xi}{\alpha}\sin\theta + \frac{\eta}{\alpha}\cos\theta\right), \\
A_2(\xi,\eta) &= \frac{1}{\alpha^2} A_1\left(\frac{\xi}{\alpha}\cos\theta - \frac{\eta}{\alpha}\sin\theta, \frac{\xi}{\alpha}\sin\theta + \frac{\eta}{\alpha}\cos\theta\right), \\
\Phi_2(\xi,\eta) &= \Phi_1\left(\frac{\xi}{\alpha}\cos\theta - \frac{\eta}{\alpha}\sin\theta, \frac{\xi}{\alpha}\sin\theta + \frac{\eta}{\alpha}\cos\theta\right) \oplus (-\xi x_0 - \eta y_0).
\end{aligned}$$

Proof. Let us suppose functions

$$\begin{aligned}
f_2(x,y) &= f_1(\alpha x\cos\theta - \alpha y\sin\theta - x_0, \alpha x\sin\theta + \alpha y\cos\theta - y_0) \\
f_3(x,y) &= f_1(\alpha x\cos\theta - \alpha y\sin\theta, \alpha x\sin\theta + \alpha y\cos\theta) \\
f_4(x,y) &= f_1(x\cos\theta - y\sin\theta, x\sin\theta + y\cos\theta)
\end{aligned}$$

Then the preceding theorems imply that

$$\begin{aligned}
F_2(\xi,\eta) &= F_3(\xi,\eta)\mathrm{e}^{-\mathrm{i}(\xi x_0+\eta y_0)} = \\
&= \frac{1}{\alpha^2} F_4\left(\frac{\xi}{\alpha}, \frac{\eta}{\alpha}\right)\mathrm{e}^{-\mathrm{i}(\xi x_0+\eta y_0)} = \\
&= \frac{1}{\alpha^2}\mathrm{e}^{-\mathrm{i}(\xi x_0+\eta y_0)} F_1\left(\frac{\xi}{\alpha}\cos\theta - \frac{\eta}{\alpha}\sin\theta, \frac{\xi}{\alpha}\sin\theta + \frac{\eta}{\alpha}\cos\theta\right).
\end{aligned}$$

The formulae for the amplitude and phase spectra are also clear consequences of the preceding theorems. □

Convention 2.47. Similarity generally involves not only rotation, scale-change and shift, but also axial symmetry. There are no special methods for registration of functions transformed by axial symmetry. From now on, I will use the notion of similarity for rotation, scale-change and shift only. The effect of a general affine transformation, which includes shift, scale-change, rotation and axial symmetry as special cases, is described by the following theorem.

Theorem 2.48 (Affine Transformation Theorem). Let function $f_1(x,y) \in \mathcal{L}(\mathbb{R}^2)$ and let $F_1(\xi,\eta)$ be its Fourier spectrum. Let us suppose function $f_2(x,y) = f_1(ax+by+c, dx+ey+g)$, where $a,b,c,d,e,g \in \mathbb{R}$ are given constants with

$$D = \begin{vmatrix} a & b \\ d & e \end{vmatrix} \neq 0.$$

The Fourier spectrum of function f_2 is [12]

$$F_2(\xi,\eta) = \frac{1}{|D|} \mathrm{e}^{\frac{\mathrm{i}}{D}[(ec-bg)\xi+(ag-cd)\eta]} F_1\left(\frac{e\xi-d\eta}{D}, \frac{-b\xi+a\eta}{D}\right).$$

Proof. The proof is taken from [12] and adapted to the definition of the Fourier transform we are using here.
Let us express the affine transformation

$$\begin{aligned} x' &= ax+by+c \\ y' &= dx+ey+g \end{aligned}$$

in matrix notation:

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} a & b \\ d & e \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} c \\ g \end{pmatrix}.$$

The Jacobian of the transformation is $|D|$, which gives the relation

$$\mathrm{d}x\,\mathrm{d}y = \frac{\mathrm{d}x'\,\mathrm{d}y'}{|D|}.$$

If the Jacobian is non-zero, the transformation can be inverted

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} a & b \\ d & e \end{pmatrix}^{-1} \begin{pmatrix} x'-c \\ y'-g \end{pmatrix}.$$

In the exponent $-\mathrm{i}(x\xi+y\eta)$, which occurs in the definition of the Fourier transform, we can rewrite the term $x\xi+y\eta$ as

$$\begin{aligned} x\xi+y\eta = (\xi\ \eta)\begin{pmatrix} x \\ y \end{pmatrix} = (\xi\ \eta)\begin{pmatrix} a & b \\ d & e \end{pmatrix}^{-1}\begin{pmatrix} x'-c \\ y'-g \end{pmatrix} = \\ = \frac{1}{D}(\xi\ \eta)\begin{pmatrix} e & -b \\ -d & a \end{pmatrix}\begin{pmatrix} x'-c \\ y'-g \end{pmatrix} = \frac{1}{D}(e\xi-d\eta \quad -b\xi+a\eta)\begin{pmatrix} x'-c \\ y'-g \end{pmatrix}. \end{aligned}$$

Plugging the last formula into the definition of the Fourier transform yields

$$F_2(\xi,\eta) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_1(ax+by+c, dx+ey+g)\mathrm{e}^{-\mathrm{i}(x\xi+y\eta)}\,\mathrm{d}x\,\mathrm{d}y =$$

$$= \mathrm{e}^{\frac{\mathrm{i}}{D}[(e\xi-d\eta)c+(-b\xi+a\eta)g]}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_1(x',y')\mathrm{e}^{-\frac{\mathrm{i}}{D}[(e\xi-d\eta)x'+(-b\xi+a\eta)y']}\frac{\mathrm{d}x'\,\mathrm{d}y'}{|D|} =$$

$$= \frac{1}{|D|}\mathrm{e}^{\frac{\mathrm{i}}{D}[(ec-bg)\xi+(ag-cd)\eta]}F_1\left(\frac{e\xi-d\eta}{D}, \frac{-b\xi+a\eta}{D}\right),$$

which was to be proved. □

Theorem 2.49. Let function $f \in \mathcal{L}(\mathbb{R}^2)$. Then [34]

(a) $$\mathcal{F}\left\{f(x,y)\right\} = 4\pi^2\mathcal{F}^{-1}\left\{f(-x,-y)\right\},$$

(b) $$\mathcal{F}^{-1}\left\{f(x,y)\right\} = \frac{1}{4\pi^2}\mathcal{F}\left\{f(-x,-y)\right\}.$$

If f is also continuous and its Fourier spectrum $F(\xi,\eta)$ is in $\mathcal{L}\left(\mathbb{R}^2\right)$, then

(c) $$\mathcal{F}\left\{\mathcal{F}\left\{f(x,y)\right\}\right\} = 4\pi^2 f(-x,-y),$$

(d) $$\mathcal{F}^{-1}\left\{\mathcal{F}^{-1}\left\{f(x,y)\right\}\right\} = \frac{1}{4\pi^2}f(-x,-y).$$

Proof.

(a)

$$4\pi^2\mathcal{F}^{-1}\left\{f(-x,-y)\right\} = 4\pi^2\frac{1}{4\pi^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(-x,-y)\mathrm{e}^{\mathrm{i}x\xi+\mathrm{i}y\eta}\,\mathrm{d}x\,\mathrm{d}y =$$

$$= \left|\begin{array}{ll} s=-x & \mathrm{d}s=-\mathrm{d}x \\ t=-y & \mathrm{d}s=-\mathrm{d}y \end{array}\right| = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(s,t)\mathrm{e}^{-\mathrm{i}s\xi-\mathrm{i}t\eta}\,\mathrm{d}s\,\mathrm{d}t = \mathcal{F}\left\{f(x,y)\right\}.$$

(b) is obtained from (a) by substituting $-x$ for x and $-y$ for y and dividing the equality by $4\pi^2$.

(c)

$$\begin{aligned}
\mathcal{F}\left\{\mathcal{F}\left\{f(x,y)\right\}\right\} &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} F(\xi,\eta)\mathrm{e}^{-\mathrm{i}(x\xi+y\eta)}\,\mathrm{d}\xi\,\mathrm{d}\eta = \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} F(\xi,\eta)\mathrm{e}^{\mathrm{i}x(-\xi)+\mathrm{i}y(-\eta)}\,\mathrm{d}\xi\,\mathrm{d}\eta = \left|\begin{array}{ll} \sigma=-\xi & \mathrm{d}\sigma=-\mathrm{d}\xi \\ \tau=-\eta & \mathrm{d}\tau=-\mathrm{d}\eta \end{array}\right| = \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} F(-\sigma,-\tau)\mathrm{e}^{\mathrm{i}x\sigma+\mathrm{i}y\tau}\,\mathrm{d}\sigma\,\mathrm{d}\tau = \\
&= 4\pi^2\mathcal{F}^{-1}\left\{\mathcal{F}\left\{f(-x,-y)\right\}\right\} = 4\pi^2 f(-x,-y),
\end{aligned}$$

the pre-last equality due to Corollary 2.44 (page 24).

(d)

$$\mathcal{F}^{-1}\left\{\mathcal{F}^{-1}\left\{f(x,y)\right\}\right\} = \frac{1}{4\pi^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left(\frac{1}{4\pi^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(s,t)\mathrm{e}^{\mathrm{i}s\xi+\mathrm{i}t\eta}\,\mathrm{d}s\,\mathrm{d}t\right)\mathrm{e}^{\mathrm{i}x\xi+\mathrm{i}y\eta}\,\mathrm{d}\xi\,\mathrm{d}\eta =$$

$$= \frac{1}{16\pi^4}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left(\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(s,t)\mathrm{e}^{-\mathrm{i}(-s)\xi-\mathrm{i}(-t)\eta}\,\mathrm{d}s\,\mathrm{d}t\right)\mathrm{e}^{\mathrm{i}x\xi+\mathrm{i}y\eta}\,\mathrm{d}\xi\,\mathrm{d}\eta =$$

$$= \left|\begin{array}{ll} u=-s & \mathrm{d}u=-\mathrm{d}s \\ v=-t & \mathrm{d}v=-\mathrm{d}t \end{array}\right| =$$

$$= \frac{1}{16\pi^4}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left(\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(-u,-v)\mathrm{e}^{-\mathrm{i}u\xi-\mathrm{i}v\eta}\,\mathrm{d}u\,\mathrm{d}v\right)\mathrm{e}^{\mathrm{i}x\xi+\mathrm{i}y\eta}\,\mathrm{d}\xi\,\mathrm{d}\eta =$$

$$= \frac{1}{4\pi^2}\mathcal{F}^{-1}\left\{\mathcal{F}\left\{f(-x,-y)\right\}\right\} = \frac{1}{4\pi^2}f(-x,-y),$$

the pre-last equality due to Corollary 2.44 (page 24).

□

Example 2.50. The Fourier transform of the centered ($\mu = 0$) two-dimensional Gauss function [35] with variance matrix $\sigma^2\mathbf{I}$ is function $\mathrm{e}^{-\frac{\sigma^2\left(\xi^2+\eta^2\right)}{2}}$, i.e. Gauss function with variance matrix $\sigma^{-2}\mathbf{I}$ without the normalizing factor

$$\mathcal{F}\left\{\frac{1}{2\pi\sigma^2}\mathrm{e}^{-\frac{x^2+y^2}{2\sigma^2}}\right\} = \mathrm{e}^{-\frac{\sigma^2\left(\xi^2+\eta^2\right)}{2}}.$$

Solution. The solution is taken from [34] and adapted for two dimensions. Let

$$f(x,y) = \frac{1}{2\pi\sigma^2}\mathrm{e}^{-\frac{x^2+y^2}{2\sigma^2}}.$$

Its Fourier transform is

$$\mathcal{F}\left\{f(x,y)\right\} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\frac{1}{2\pi\sigma^2}\mathrm{e}^{-\frac{x^2+y^2}{2\sigma^2}}\mathrm{e}^{-\mathrm{i}(\xi x+\eta y)}\,\mathrm{d}x\,\mathrm{d}y =$$

$$= \underbrace{\frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{\infty}\mathrm{e}^{-\frac{x^2}{2\sigma^2}}\mathrm{e}^{-\mathrm{i}\xi x}\,\mathrm{d}x}_{\text{denote by } G(\xi)} \cdot \underbrace{\frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{\infty}\mathrm{e}^{-\frac{y^2}{2\sigma^2}}\mathrm{e}^{-\mathrm{i}\eta y}\,\mathrm{d}y}_{\text{denote by } G(\eta)}$$

Since G is a positive function, we can split the integral into two

$$G(\xi) = \frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{\infty}\mathrm{e}^{-\frac{x^2}{2\sigma^2}}\cos\xi x\,\mathrm{d}x + \underbrace{\frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{\infty}\mathrm{e}^{-\frac{x^2}{2\sigma^2}}\mathrm{i}\sin\xi x\,\mathrm{d}x}_{=0} =$$

$$= \frac{1}{\sigma\sqrt{2\pi}}2\int_{0}^{\infty}\mathrm{e}^{-\frac{x^2}{2\sigma^2}}\cos\xi x\,\mathrm{d}x = \frac{2}{\sigma\sqrt{2\pi}}\int_{0}^{\infty}\mathrm{e}^{-\frac{x^2}{2\sigma^2}}\sum_{n=0}^{\infty}\frac{(-1)^n}{(2n)!}(\xi x)^{2n}\,\mathrm{d}x.$$

By Beppo Levi's Theorem [9],

$$G(\xi) = \frac{2}{\sigma\sqrt{2\pi}} \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} \xi^{2n} \underbrace{\int_0^{\infty} x^{2n} \mathrm{e}^{-\frac{x^2}{2\sigma^2}} \,\mathrm{d}x}_{\text{denote by } I_n}.$$

By partial integration we obtain I_n as

$$\begin{aligned} I_n &= \left| \begin{array}{cccccc} u & = & \mathrm{e}^{-\frac{x^2}{2\sigma^2}} & u' & = & -\frac{2x}{2\sigma^2} \cdot \mathrm{e}^{-\frac{x^2}{2\sigma^2}} \\ v' & = & x^{2n} & v & = & \frac{x^{2n+1}}{2n+1} \end{array} \right| = \\ &= \left[\mathrm{e}^{-\frac{x^2}{2\sigma^2}} \cdot \frac{x^{2n+1}}{2n+1} \right]_{x=0}^{\infty} - \int_0^{\infty} \frac{x^{2n+1}}{2n+1} \left(-\frac{x}{\sigma^2} \right) \mathrm{e}^{-\frac{x^2}{2\sigma^2}} \,\mathrm{d}x = \\ &= 0 + \frac{1}{\sigma^2(2n+1)} \int_0^{\infty} x^{2(n+1)} \mathrm{e}^{-\frac{x^2}{2\sigma^2}} \,\mathrm{d}x = \frac{1}{\sigma^2(2n+1)} I_{n+1}. \end{aligned}$$

For $n > 0$

$$I_n = \sigma^2(2n-1)I_{n-1} = \left(\sigma^2\right)^2 (2n-1)(2n-3)I_{n-2} = (2n-1)!! \left(\sigma^2\right)^n I_0.$$

$$I_0 = \int_0^{\infty} \mathrm{e}^{-\frac{x^2}{2\sigma^2}} \,\mathrm{d}x = \frac{\sigma\sqrt{2\pi}}{2}$$

$$(2n-1)!! = (2n-1)(2n-3) \cdot \cdots \cdot 3 \cdot 1 = \frac{(2n-1)!}{(2n-2)(2n-4) \cdot \cdots \cdot 4 \cdot 2} = \frac{(2n-1)!}{(n-1)!2^{n-1}}$$

$$I_n = \frac{(2n-1)!}{(n-1)!2^{n-1}} \sigma^{2n} \cdot \frac{\sigma\sqrt{2\pi}}{2}$$

$$\begin{aligned} G(\xi) &= \frac{2}{\sigma\sqrt{2\pi}} \left[\frac{\sigma\sqrt{2\pi}}{2} + \sum_{n=1}^{\infty} \frac{(-1)^n}{(2n)!} \xi^{2n} \frac{(2n-1)!}{(n-1)!2^{n-1}} \sigma^{2n} \cdot \frac{\sigma\sqrt{2\pi}}{2} \right] = \\ &= 1 + \sum_{n=1}^{\infty} \xi^{2n} \sigma^{2n} (-1)^n \frac{1}{n!2^n} = \sum_{n=0}^{\infty} \frac{\xi^{2n} \sigma^{2n} (-1)^n}{n!2^n} = \sum_{n=0}^{\infty} \frac{1}{n!} \left[-\frac{(\xi\sigma)^2}{2} \right]^n = \mathrm{e}^{-\frac{\xi^2\sigma^2}{2}} \end{aligned}$$

$$\mathcal{F}\{f(x,y)\} = G(\xi)G(\eta) = \mathrm{e}^{-\frac{\xi^2\sigma^2}{2}} \cdot \mathrm{e}^{-\frac{\eta^2\sigma^2}{2}} = \mathrm{e}^{-\frac{\sigma^2\left(\xi^2+\eta^2\right)}{2}}$$

□

In the following examples, the Fourier transforms of functions which are not in $\mathcal{L}(\mathbb{R}^2)$ are presented. The functions will used for proves of the phase-correlation function properties. In these cases, the derivation involves the theory of distributions.

Example 2.51. The Fourier transform of the Dirac distribution shifted by $(x_0, y_0) \in \mathbb{R}^2$ is $\mathrm{e}^{-\mathrm{i}(\xi x_0 + \eta y_0)}$

$$\mathcal{F}\{\delta(x - x_0, y - y_0)\} = \mathrm{e}^{-\mathrm{i}(\xi x_0 + \eta y_0)}.$$

Solution. The derivation can be found in [42]. □

Example 2.52. The Fourier transform of exponential function $\mathrm{e}^{\mathrm{i}(ax+by)}$, $a, b \in \mathbb{R}$ is the Dirac distribution shifted by (a, b)

$$\mathcal{F}\left\{\mathrm{e}^{\mathrm{i}(ax+by)}\right\} = \delta(x-a, y-b).$$

Solution. The derivation can be found in [42]. □

Theorem 2.53. Let $f(x, y) \in \mathcal{L}(\mathbb{R}^2)$ and let $F(\xi, \eta)$ be its Fourier spectrum. The Fourier spectrum of the complex conjugate of function f is the complex conjugate of its Fourier spectra with reversed axes [34]

$$\mathcal{F}\left\{f^*(x, y)\right\} = F^*(-\xi, -\eta).$$

Proof. Proof taken from [34].

$$\begin{aligned}\mathcal{F}\left\{f^*(x,y)\right\} &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f^*(x,y)\mathrm{e}^{-\mathrm{i}(x\xi+y\eta)}\,\mathrm{d}x\,\mathrm{d}y = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f^*(x,y)\mathrm{e}^{\mathrm{i}(-x\xi-y\eta)}\,\mathrm{d}x\,\mathrm{d}y = \\ &= \left(\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\mathrm{e}^{-\mathrm{i}[x(-\xi)+y(-\eta)]}\,\mathrm{d}x\,\mathrm{d}y\right)^* = F^*(-\xi,-\eta),\end{aligned}$$

where the third equality holds because for $a \in \mathbb{R}$ is $\mathrm{e}^{\mathrm{i}a} = \cos a + \mathrm{i}\sin a$, $\mathrm{e}^{-\mathrm{i}a} = \cos a + \mathrm{i}\sin(-a) = \cos a - \mathrm{i}\sin a$. Hence $\mathrm{e}^{\mathrm{i}a} = \left(\mathrm{e}^{-\mathrm{i}a}\right)^*$. □

Theorem 2.54. Let $f(x, y) \in \mathcal{L}(\mathbb{R}^2)$ and let $F(\xi, \eta) \in \mathcal{L}(\mathbb{R}^2)$ be its Fourier spectrum. Let f be continuous. Then the inverse Fourier transform of the complex conjugate of spectrum F is the complex conjugate of function f with reversed axes, i.e. in every point where f is continuous it holds

$$\mathcal{F}^{-1}\left\{F^*(\xi, \eta)\right\} = f^*(-x, -y).$$

Proof. Proof taken from [34].

$$\begin{aligned}\mathcal{F}^{-1}\left\{F^*(\xi,\eta)\right\} &= \frac{1}{4\pi^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} F^*(\xi,\eta)\mathrm{e}^{\mathrm{i}(\xi x+\eta y)}\,\mathrm{d}\xi\,\mathrm{d}\eta = \\ &= \frac{1}{4\pi^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} F^*(\xi,\eta)\mathrm{e}^{-\mathrm{i}(-\xi x-\eta y)}\,\mathrm{d}\xi\,\mathrm{d}\eta = \\ &= \left(\frac{1}{4\pi^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} F(\xi,\eta)\mathrm{e}^{\mathrm{i}[\xi(-x)+\eta(-y)]}\,\mathrm{d}\xi\,\mathrm{d}\eta\right)^* = f^*(-x,-y)\end{aligned}$$

□

Theorem 2.55. Let $f(x, y) \in \mathcal{L}(\mathbb{R}^2)$ and let $F(\xi, \eta) \in \mathcal{L}(\mathbb{R}^2)$ be its Fourier spectrum. Let f be continuous. f is a real function, i.e. $f(x, y) = f^*(x, y)\ \forall (x, y) \in \mathbb{R}^2$, if and only if $F(\xi, \eta) = F^*(-\xi, -\eta)$.

Proof. Proof taken from [34].

(a) Let us suppose that f is a real function. Then Theorem 2.53 implies that

$$F(\xi,\eta) = \mathcal{F}\{f(x,y)\} = \mathcal{F}\{f^*(x,y)\} = F^*(-\xi,-\eta).$$

(b) Let us suppose that $F(\xi,\eta) = F^*(-\xi,-\eta)$. Then Theorem 2.54 implies that

$$f(x,y) = \mathcal{F}^{-1}\{F(\xi,\eta)\} = \mathcal{F}^{-1}\{F^*(-\xi,-\eta)\} = f^*(x,y).$$

□

Corollary 2.56. Let $f(x,y) \in \mathcal{L}(\mathbb{R}^2)$ be a real function with a Fourier spectrum $F(\xi,\eta)$. Then

$$A(\xi,\eta) = A(-\xi,-\eta).$$

Proof. Consequence of Theorem 2.53. We can also use the first part of the proof of Theorem 2.55, which does not require that $F(\xi,\eta) \in \mathcal{L}(\mathbb{R}^2)$. □

Corollary 2.57. Let $f(x,y) \in \mathcal{L}(\mathbb{R}^2)$ be a real continuous function with a Fourier spectrum $F(\xi,\eta) \in \mathcal{L}(\mathbb{R}^2)$. Let $G(\xi,\eta)$ be a bounded function $\mathbb{R}^2 \to \mathbb{R}$ such that $G(\xi,\eta) = G(-\xi,-\eta)$. Then

$$\mathcal{F}^{-1}\{F(\xi,\eta)\cdot G(\xi,\eta)\}$$

is real.

Proof. According to Theorem 2.55, if f is real, then

$$F(\xi,\eta) = F^*(-\xi,-\eta).$$

Multiplying the equality by G, we obtain

$$F(\xi,\eta)\cdot G(\xi,\eta) = F^*(-\xi,-\eta)\cdot G(-\xi,-\eta) = (F(-\xi,-\eta)\cdot G(-\xi,-\eta))^*.$$

Since G is bounded, there is no doubt about existence of the inverse Fourier transform. Then again according to Theorem 2.55,

$$\mathcal{F}^{-1}\{F(\xi,\eta)\cdot G(\xi,\eta)\}$$

is real. □

2.5 Convolution

Definition 2.58 (Convolution). Let functions $f_1(x,y), f_2(x,y) \in \mathcal{L}(\mathbb{R}^2)$. The *convolution* [8] $f_1 * f_2$ of functions f_1, f_2 is a function

$$f(x,y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_1(s,t) f_2(x-s,y-t)\,\mathrm{d}s\,\mathrm{d}t.$$

Theorem 2.59. Let functions $f_1(x,y), f_2(x,y) \in \mathcal{L}(\mathbb{R}^2)$. Then $f_1 * f_2 \in \mathcal{L}(\mathbb{R}^2)$.

Proof. We start by proving that $f_1(x,y)\cdot f_2(u,v)\in\mathcal{L}\,(\mathbb{R}^4)$, i.e. (a proof from [10])

$$\int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty}|f_1(x,y)f_2(u,v)|\,\mathrm{d}u\,\mathrm{d}v\,\mathrm{d}x\,\mathrm{d}y<\infty.$$

The definitions of the quadruple improper integral and the $\mathcal{L}(\mathbb{R}^4)$ space are only technical. They are similar to the definitions in the two-dimensional case and are can be found in Appendix D on page 99.

$$\begin{aligned}&\int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty}|f_1(x,y)f_2(u,v)|\,\mathrm{d}u\,\mathrm{d}v\,\mathrm{d}x\,\mathrm{d}y=\\&=\int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty}|f_1(x,y)|\cdot|f_2(u,v)|\,\mathrm{d}u\,\mathrm{d}v\,\mathrm{d}x\,\mathrm{d}y=\\&=\int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty}|f_1(x,y)|\underbrace{\left(\int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty}|f_2(u,v)|\,\mathrm{d}u\,\mathrm{d}v\right)}_{\text{denote by }I}\mathrm{d}x\,\mathrm{d}y\end{aligned}$$

$0\le I<\infty$ since $f_2\in\mathcal{L}\,(\mathbb{R}^2)$. Therefore using Fubini's Theorem [9],

$$\begin{aligned}&\int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty}|f_1(x,y)|\left(\int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty}|f_2(u,v)|\,\mathrm{d}u\,\mathrm{d}v\right)\mathrm{d}x\,\mathrm{d}y=\\&=\int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty}|f_1(x,y)|\,\mathrm{d}x\,\mathrm{d}y\int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty}|f_2(u,v)|\,\mathrm{d}u\,\mathrm{d}v<\infty.\end{aligned}$$

Thus $f_1(x,y)\cdot f_2(u,v)\in\mathcal{L}\,(\mathbb{R}^4)$. By making the change of variables $u=p-s,x=s$, $v=q-t,y=t$ we obtain

$$\int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty}f_1(x,y)f_2(u,v)\,\mathrm{d}u\,\mathrm{d}v\,\mathrm{d}x\,\mathrm{d}y=\int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty}f_1(s,t)f_2(p-s,q-t)\,\mathrm{d}s\,\mathrm{d}t\,\mathrm{d}p\,\mathrm{d}q,$$

which says that function

$$(f_1*f_2)(p,q)=\int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty}f_1(p-s,q-t)f_2(s,t)\,\mathrm{d}s\,\mathrm{d}t$$

belongs to $\mathcal{L}\,(\mathbb{R}^2)$. □

Theorem 2.60. Let functions $f_1(x,y),f_2(x,y)\in\mathcal{L}\,(\mathbb{R}^2)$ with Fourier spectra $F_1(\xi,\eta)$, $F_2(\xi,\eta)$. Then

$$\mathcal{F}\left\{f_1(x,y)*f_2(x,y)\right\}=F_1(\xi,\eta)\cdot F_2(\xi,\eta).$$

Proof. Proof taken from [8].

$$\mathcal{F}\{f_1 * f_2\} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left(\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_1(s,t)f_2(x-s,y-t)\,\mathrm{d}s\,\mathrm{d}t\right)\mathrm{e}^{-\mathrm{i}(x\xi+y\eta)}\,\mathrm{d}x\,\mathrm{d}y =$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left(\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_1(s,t)f_2(x-s,y-t)\mathrm{e}^{-\mathrm{i}(x\xi+y\eta)}\,\mathrm{d}x\,\mathrm{d}y\right)\mathrm{d}s\,\mathrm{d}t =$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_1(s,t)\left(\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_2(x-s,y-t)\mathrm{e}^{-\mathrm{i}(x\xi+y\eta)}\,\mathrm{d}x\,\mathrm{d}y\right)\mathrm{d}s\,\mathrm{d}t =$$

$$= \left|\begin{array}{llll} p = x-s & x = s+p & \mathrm{d}x = & \mathrm{d}p \\ q = y-t & y = t+q & \mathrm{d}y = & \mathrm{d}q \end{array}\right| =$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_1(s,t)\left(\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_2(p,q)\mathrm{e}^{-\mathrm{i}\xi(p+s)-\mathrm{i}\eta(q+t)}\,\mathrm{d}p\,\mathrm{d}q\right)\mathrm{d}s\,\mathrm{d}t =$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_1(s,t)\left(\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_2(p,q)\mathrm{e}^{-\mathrm{i}(\xi p+\eta q)}\mathrm{e}^{-\mathrm{i}(\xi s+\eta t)}\,\mathrm{d}p\,\mathrm{d}q\right)\mathrm{d}s\,\mathrm{d}t =$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_1(s,t)\mathrm{e}^{-\mathrm{i}(\xi s+\eta t)}\,\mathrm{d}s\,\mathrm{d}t \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_2(p,q)\mathrm{e}^{-\mathrm{i}(\xi p+\eta q)}\,\mathrm{d}p\,\mathrm{d}q = F_1(\xi,\eta)\cdot F_2(\xi,\eta)$$

□

Theorem 2.61. Let functions $f_1(x,y), f_2(x,y) \in \mathcal{L}(\mathbb{R}^2)$ with Fourier spectra $F_1(\xi,\eta)$, $F_2(\xi,\eta) \in \mathcal{L}(\mathbb{R}^2)$. Let f_1, f_2 be continuous. Then

$$\mathcal{F}\{f_1(x,y)\cdot f_2(x,y)\} = \frac{1}{4\pi^2}F_1(\xi,\eta) * F_2(\xi,\eta).$$

Proof. Proof taken from [32] adapted for two dimensions.

$$\mathcal{F}\{f_1(x,y)\cdot f_2(x,y)\} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_1(x,y)f_2(x,y)\mathrm{e}^{-\mathrm{i}(x\xi+y\eta)}\,\mathrm{d}x\,\mathrm{d}y =$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left(\frac{1}{4\pi^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} F_1(\sigma,\tau)\mathrm{e}^{\mathrm{i}(x\sigma+y\tau)}\,\mathrm{d}\sigma\,\mathrm{d}\tau\right) f_2(x,y)\mathrm{e}^{-\mathrm{i}(x\xi+y\eta)}\,\mathrm{d}x\,\mathrm{d}y =$$

$$= \frac{1}{4\pi^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} F_1(\sigma,\tau)\left(\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_2(x,y)\mathrm{e}^{-\mathrm{i}x(\xi-\sigma)-\mathrm{i}y(\eta-\tau)}\,\mathrm{d}x\,\mathrm{d}y\right)\mathrm{d}\sigma\,\mathrm{d}\tau =$$

$$= \frac{1}{4\pi^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} F_1(\sigma,\tau)F_2(\xi-\sigma,\eta-\tau)\,\mathrm{d}\sigma\,\mathrm{d}\tau = \frac{1}{4\pi^2}F_1(\xi,\eta) * F_2(\xi,\eta)$$

□

2.6 Cross correlation, phase correlation

Definition 2.62 (Cross-power spectrum, normalized cross-power spectrum, semi-normalized cross-power spectrum). Let functions $f_1(x,y), f_2(x,y) \in \mathcal{L}(\mathbb{R}^2)$ have Fourier spectra $F_1(\xi,\eta)$, $F_2(\xi,\eta)$. The *cross-power spectrum* [13] of functions f_1, f_2 is function $C_{f_1,f_2}(\xi,\eta) : \mathbb{R}^2 \to \mathbb{C}$ defined as

$$C_{f_1,f_2}(\xi,\eta) = F_1(\xi,\eta) \cdot F_2^*(\xi,\eta).$$

The *normalized cross-power spectrum* [13] is function $Z_{f_1,f_2}(\xi,\eta) : \mathbb{R}^2 \to \mathbb{C}$ defined as

$$Z_{f_1,f_2}(\xi,\eta) = \frac{F_1(\xi,\eta) \cdot F_2^*(\xi,\eta)}{|F_1(\xi,\eta) \cdot F_2(\xi,\eta)|},$$

and if $p, q \in \mathbb{R}^+$ are given constants, the *semi-normalized cross-power spectrum* is function $Z^{p,q}_{f_1,f_2}(\xi,\eta) : \mathbb{R}^2 \to \mathbb{C}$ defined as

$$Z^{p,q}_{f_1,f_2}(\xi,\eta) = \frac{F_1(\xi,\eta) \cdot F_2^*(\xi,\eta)}{(|F_1(\xi,\eta)| + p) \cdot (|F_2(\xi,\eta)| + q)}.$$

Definition 2.63 (Cross-correlation function, phase-correlation function, semi-phase correlation function). Let functions $f_1(x,y), f_2(x,y) \in \mathcal{L}(\mathbb{R}^2)$ have Fourier spectra $F_1(\xi,\eta)$, $F_2(\xi,\eta)$. Function $Q_{f_1,f_2}(x,y) : \mathbb{R}^2 \to \mathbb{C}$ defined as

$$Q_{f_1,f_2}(x,y) = \mathcal{F}^{-1}\left\{C_{f_1,f_2}(\xi,\eta)\right\} = \mathcal{F}^{-1}\left\{F_1(\xi,\eta) \cdot F_2^*(\xi,\eta)\right\}$$

is called the *cross-correlation function* of functions f_1, f_2. Function $P_{f_1,f_2}(x,y) : \mathbb{R}^2 \to \mathbb{C}$ defined as

$$P_{f_1,f_2}(x,y) = \mathcal{F}^{-1}\left\{Z_{f_1,f_2}(\xi,\eta)\right\} = \mathcal{F}^{-1}\left\{\frac{F_1(\xi,\eta) \cdot F_2^*(\xi,\eta)}{|F_1(\xi,\eta) \cdot F_2(\xi,\eta)|}\right\} \tag{2.6}$$

is called the *phase-correlation function* of functions f_1, f_2. Function $P^{p,q}_{f_1,f_2}(x,y) : \mathbb{R}^2 \to \mathbb{C}$ defined as

$$P^{p,q}_{f_1,f_2}(x,y) = \mathcal{F}^{-1}\left\{Z^{p,q}_{f_1,f_2}(\xi,\eta)\right\} = \mathcal{F}^{-1}\left\{\frac{F_1(\xi,\eta) \cdot F_2^*(\xi,\eta)}{(|F_1(\xi,\eta)| + p) \cdot (|F_2(\xi,\eta)| + q)}\right\}$$

is called the *semi-phase correlation function* of functions f_1, f_2.

Theorem 2.64 (Cross-correlation function of real functions). Let functions $f_1(x, y)$, $f_2(x,y) \in \mathcal{L}(\mathbb{R}^2)$ be continuous real ($f_1^* = f_1, f_2^* = f_2$) and with Fourier spectra $F_1(\xi,\eta)$, $F_2(\xi,\eta) \in \mathcal{L}(\mathbb{R}^2)$. Then the cross-correlation function of these functions is real.

Proof. Using Theorems 2.37 (page 20), 2.54 (page 30), and 2.60 (page 32), we can compute the cross-correlation function of f_1 and f_2 as

$$\begin{aligned} Q_{f_1,f_2}(x,y) = \mathcal{F}^{-1}\left\{F_1(\xi,\eta) \cdot F_2^*(\xi,\eta)\right\} = \mathcal{F}^{-1}\left\{\mathcal{F}\left\{f_1(x,y)\right\} \cdot \mathcal{F}\left\{f_2^*(-x,-y)\right\}\right\} = \\ = f_1(x,y) * f_2^*(-x,-y) = f_1(x,y) * f_2(-x,-y), \end{aligned}$$

which is a real function. □

Remark 2.65. In [34] and in one-dimensional case in [32], we can find the definition of the *cross-correlation* of functions $f_1(x,y), f_2(x,y)$ as

$$f_1(x,y) \star f_2(x,y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_1^*(s,t) f_2(s+x,t+y)\,\mathrm{d}x\,\mathrm{d}y.$$

The cross-correlation can be transformed to convolution [34]

$$f_1(x,y) \star f_2(x,y) = f_1^*(-x,-y) * f_2(x,y).$$

Therefore, using the proof of Theorem 2.64 and commutativity of convolution [15]

$$Q_{f_1,f_2}(x,y) = f_1(x,y) * f_2^*(-x,-y) = f_2^*(-x,-y) * f_1(x,y) = f_2(x,y) \star f_1(x,y),$$

which gives the connection between the cross-correlation used in Fourier analysis [32] and the cross-correlation function Q_{f_1,f_2} defined in Definition 2.63 and used in image registration.

Theorem 2.66. Let functions $f_1(x,y),\ f_2(x,y) \in \mathcal{L}(\mathbb{R}^2)$ be real with Fourier spectra $F_1(\xi,\eta), F_2(\xi,\eta)$ and amplitude spectra $A_1(\xi,\eta), A_2(\xi,\eta)$. Let $p,q \in \mathbb{R}_0^+$, let $\varepsilon \in \mathbb{R}^+$. If

$$A_1(\xi,\eta), A_2(\xi,\eta) > \varepsilon\ \forall(\xi,\eta) \in \mathbb{R}^2 \quad \text{or} \quad p,q > 0$$

then the semi-phase correlation function of functions f_1, f_2 with parameters p,q is a real function, i.e.

$$\mathcal{F}^{-1}\left\{\frac{F_1(\xi,\eta)\cdot F_2^*(\xi,\eta)}{(A_1(\xi,\eta)+p)\cdot(A_2(\xi,\eta)+q)}\right\}$$

is a real function. Especially, if $A_1(\xi,\eta), A_2(\xi,\eta) > \varepsilon\ \forall(\xi,\eta) \in \mathbb{R}^2$, the phase-correlation function of functions f_1, f_2 is real.

Proof. Amplitude spectra are real functions and from Corollary 2.56 (page 31) we have that for real functions

$$A_k(\xi,\eta) = A_k(-\xi,-\eta) \ \text{ for } k = 1,2.$$

Furthermore, if

$$A(\xi,\eta) = (A_1(\xi,\eta)+p)\cdot(A_2(\xi,\eta)+q),$$

we have $A(\xi,\eta) = A(-\xi,-\eta)$ and A is again real. Now function $\frac{1}{A(\xi,\eta)}$ has all properties of function G in Corollary 2.57 (page 31). It is a real bounded function with $A(\xi,\eta) = A(-\xi,-\eta)$. Function $F_1(\xi,\eta)\cdot F_2^*(\xi,\eta)$ has the properties of function F in Corollary 2.57 due to the fact that the cross-correlation function of two real functions is real (Thm. 2.64). Therefore,

$$\mathcal{F}^{-1}\left\{\frac{F_1(\xi,\eta)\cdot F_2^*(\xi,\eta)}{A(\xi,\eta)}\right\} = \mathcal{F}^{-1}\left\{\frac{F_1(\xi,\eta)\cdot F_2^*(\xi,\eta)}{(A_1(\xi,\eta)+p)\cdot(A_2(\xi,\eta)+q)}\right\}$$

is real. If $A_1(\xi,\eta), A_2(\xi,\eta) > \varepsilon > 0\ \forall(\xi,\eta) \in \mathbb{R}^2$, we can set $p = q = 0$ and conclude that the phase-correlation function of real functions is real. □

Theorem 2.67 (Phase-correlation function of shifted functions). Let $f_1(x,y) \in \mathcal{L}(\mathbb{R}^2)$ and let $F_1(\xi,\eta)$ be its Fourier spectrum. Let us suppose function $f_2(x,y) = f_1(x-x_0, y-y_0)$, where $x_0, y_0 \in \mathbb{R}$ are given numbers. Let $F_2(\xi,\eta)$ be the Fourier spectrum of function f_2. Then the phase-correlation function of functions f_1, f_2 is the Dirac distribution shifted by $(-x_0, -y_0)$ [13]

$$P_{f_1,f_2}(x,y) = \delta(x+x_0, y+y_0).$$

Proof. The Shift Theorem (Theorem 2.42 on page 22) implies that

$$Z_{f_1,f_2}(\xi,\eta) = \frac{F_1(\xi,\eta)\cdot F_1^*(\xi,\eta)\left(\mathrm{e}^{-\mathrm{i}(\xi x_0+\eta y_0)}\right)^*}{|F_1(\xi,\eta)\cdot F_1(\xi,\eta)\mathrm{e}^{-\mathrm{i}(\xi x_0+\eta y_0)}|} = \mathrm{e}^{\mathrm{i}(\xi x_0+\eta y_0)}.$$

According to Example 2.51 on page 29

$$\mathcal{F}\left\{\delta(x-x_0, y-y_0)\right\} = \mathrm{e}^{-\mathrm{i}(\xi x_0+\eta y_0)}.$$

Therefore

$$P_{f_1,f_2}(x,y) = \mathcal{F}^{-1}\left\{\mathrm{e}^{\mathrm{i}(\xi x_0+\eta y_0)}\right\} = \mathcal{F}^{-1}\left\{\mathrm{e}^{-\mathrm{i}\xi(-x_0)-\mathrm{i}\eta(-y_0)}\right\} = \delta(x+x_0, y+y_0).$$

□

[13] describes the derivation of the formula (2.6) on page 34 for the phase-correlation function. If functions f_1, f_2 are mutually shifted, their Fourier spectra differ only in the phase (see the Shift Theorem 2.42, page 22)

$$F_2(\xi,\eta) = F_1(\xi,\eta)\mathrm{e}^{-\mathrm{i}(\xi x_0+\eta y_0)}.$$

If we compute the cross-power spectrum of such images

$$C_{f_1,f_2}(\xi,\eta) = F_1(\xi,\eta)F_1^*(\xi,\eta)\mathrm{e}^{\mathrm{i}(\xi x_0+\eta y_0)},$$

we are now interested in its phase. Normalizing the cross-power spectrum, we extract the phase

$$Z_{f_1,f_2}(\xi,\eta) = \frac{F_1(\xi,\eta)\cdot F_2^*(\xi,\eta)}{|F_1(\xi,\eta)\cdot F_2(\xi,\eta)|} = \frac{F_1(\xi,\eta)\cdot F_1^*(\xi,\eta)\left(\mathrm{e}^{-\mathrm{i}(\xi x_0+\eta y_0)}\right)^*}{|F_1(\xi,\eta)\cdot F_1(\xi,\eta)\mathrm{e}^{-\mathrm{i}(\xi x_0+\eta y_0)}|} = \mathrm{e}^{\mathrm{i}(\xi x_0+\eta y_0)}.$$

Applying the inverse Fourier transform, we obtain the Dirac distribution shifted by $(-x_0, -y_0)$.

Theorem 2.67 shows a simple method for finding the mutual shift between two functions. The phase-correlation function of two mutually shifted functions is a function which is non-zero in one single point $(-x_0, -y_0)$. Numbers x_0, y_0 give the vector by which the second function is shifted in respect to the first one and therefore if function f_2 is shifted back by $(-x_0, -y_0)$, we obtain an identical function with function f_1. Modification for similarly transformed functions (shift, rotation, scale change) will be discussed in Chapter 4.

Chapter 3

The discrete Fourier transform

3.1 The discrete Fourier transform and the inverse discrete Fourier transform

Definition 3.1 (Discrete Fourier transform). Let $f(x,y)$ be a function $\{0,1,\ldots,N-1\}\times\{0,1,\ldots,N-1\}=\{0,1,\ldots,N-1\}^2\to\mathbb{C}, N\in\mathbb{N}$. The *discrete Fourier transform* [8] of function $f(x,y)$ is function $\mathcal{D}\{f\}(\xi,\eta)=F(\xi,\eta):\{0,1,\ldots,N-1\}^2\to\mathbb{C}$ defined as

$$F(\xi,\eta)=\sum_{x=0}^{N-1}\sum_{y=0}^{N-1}f(x,y)\mathrm{e}^{-\frac{2\pi\mathrm{i}}{N}(x\xi+y\eta)}. \tag{3.1}$$

Function F is also called the *Fourier spectrum* of function f.

Definition 3.2 (Inverse discrete Fourier transform). Let $f(x,y)$ be a function $\{0,1,\ldots,N-1\}^2\to\mathbb{C}, N\in\mathbb{N}$ and let $F(\xi,\eta)$ be its discrete Fourier transform. The *inverse discrete Fourier transform* [8] of function $F(\xi,\eta)$ is function $\mathcal{D}^{-1}\{F\}(x,y):\{0,1,\ldots,N-1\}^2\to\mathbb{C}$ defined as

$$\mathcal{D}^{-1}\{F\}(x,y)=\frac{1}{N^2}\sum_{\xi=0}^{N-1}\sum_{\eta=0}^{N-1}F(\xi,\eta)\mathrm{e}^{\frac{2\pi\mathrm{i}}{N}(x\xi+y\eta)}. \tag{3.2}$$

Convention 3.3. If it makes no confusion, we will omit the word "discrete" knowing what kind of functions we are dealing with.

Theorem 3.4 (Fourier Inversion Theorem). Let $f(x,y)$ be a function $\{0,1,\ldots,N-1\}^2\to\mathbb{C}, N\in\mathbb{N}$ and let $F(\xi,\eta)$ be its discrete Fourier transform. Then the inverse discrete Fourier transform of function $F(\xi,\eta)$ is function $f(x,y)$, i.e.

$$\mathcal{D}^{-1}\{\mathcal{D}\{f(x,y)\}\}=f(x,y).$$

Proof. The proof is a generalized version of a one-dimensional proof in [41].

$$\begin{aligned}
\mathcal{D}^{-1}\left\{\mathcal{D}\left\{f(x,y)\right\}\right\} &= \frac{1}{N^2}\sum_{\xi=0}^{N-1}\sum_{\eta=0}^{N-1} F(\xi,\eta)\mathrm{e}^{\frac{2\pi\mathrm{i}}{N}(x\xi+y\eta)} = \\
&= \frac{1}{N^2}\sum_{\xi=0}^{N-1}\sum_{\eta=0}^{N-1}\sum_{s=0}^{N-1}\sum_{t=0}^{N-1} f(s,t)\mathrm{e}^{-\frac{2\pi\mathrm{i}}{N}(s\xi+t\eta)}\mathrm{e}^{\frac{2\pi\mathrm{i}}{N}(x\xi+y\eta)} = \\
&= \frac{1}{N^2}\sum_{s=0}^{N-1}\sum_{t=0}^{N-1} f(s,t)\sum_{\xi=0}^{N-1}\sum_{\eta=0}^{N-1}\mathrm{e}^{\xi\frac{2\pi\mathrm{i}}{N}(x-s)}\mathrm{e}^{\eta\frac{2\pi\mathrm{i}}{N}(y-t)} = \\
&= \frac{1}{N^2}\sum_{s=0}^{N-1}\sum_{t=0}^{N-1} f(s,t)\underbrace{\left[\sum_{\xi=0}^{N-1}\left(\mathrm{e}^{\frac{2\pi\mathrm{i}}{N}(x-s)}\right)^{\xi}\right]}_{\text{denote by } g(s)}\underbrace{\left[\sum_{\eta=0}^{N-1}\left(\mathrm{e}^{\frac{2\pi\mathrm{i}}{N}(y-t)}\right)^{\eta}\right]}_{\text{denote by } g(t)}
\end{aligned}$$

$g(s)$ is a finite geometrical series, therefore we can compute its sum. If $\mathrm{e}^{\frac{2\pi\mathrm{i}}{N}(x-s)} = 1$, i.e. $x = s$, then $g(s) = N$. Otherwise, $x - s \in \mathbb{Z} - \{0\}$ and

$$g(s) = \frac{1-\left(\mathrm{e}^{\frac{2\pi\mathrm{i}}{N}(x-s)}\right)^{N}}{1-\mathrm{e}^{\frac{2\pi\mathrm{i}}{N}(x-s)}} = \frac{1-\mathrm{e}^{2\pi\mathrm{i}(x-s)}}{1-\mathrm{e}^{\frac{2\pi\mathrm{i}}{N}(x-s)}} = \frac{1-1}{1-\mathrm{e}^{\frac{2\pi\mathrm{i}}{N}(x-s)}} = 0.$$

Similarly,

$$g(t) = \begin{cases} N & \text{if } y = t \\ 0 & \text{else.} \end{cases}$$

Therefore,

$$\mathcal{D}^{-1}\left\{\mathcal{D}\left\{f(x,y)\right\}\right\} = \frac{1}{N^2} f(x,y)\cdot N\cdot N = f(x,y).$$

□

Remark 3.5. On the contrary to the Fourier transform, the discrete Fourier transform always exists due to the fact that the summation is over a finite number of points.

Equations (3.1) and (3.2) define $F(\xi,\eta)$ and $\mathcal{D}^{-1}\left\{F\right\}(x,y)$ for every (ξ,η) and $(x,y) \in \mathbb{Z}^2$. Therefore, we can define the periodization of the Fourier spectrum, $\widetilde{F}(\xi,\eta)$, and the periodization of the original function, $\widetilde{f}(x,y)$. We define here the periodization by means of the discrete Fourier transform. This approach is used in one dimension in [15]. There are other approaches, e.g. [8] uses modulo arithmetics in formulae where the argument of the functions would be out of $\{0,1,\ldots,N-1\}^2$.

Definition 3.6 (Periodization of function and its Fourier spectrum). Let $f(x,y)$ be a function $\{0,1,\ldots,N-1\}^2 \to \mathbb{C}, N \in \mathbb{N}$ and let $F(\xi,\eta)$ be its Fourier spectrum. The *periodization of the Fourier spectrum* F is function $\widetilde{F}(\xi,\eta) : \mathbb{Z}^2 \to \mathbb{C}$ defined as

$$\widetilde{F}(\xi,\eta) = \sum_{x=0}^{N-1}\sum_{y=0}^{N-1} f(x,y)\mathrm{e}^{-\frac{2\pi\mathrm{i}}{N}(x\xi+y\eta)}.$$

The *periodization of function* f is function $\widetilde{f}(x,y) : \mathbb{Z}^2 \to \mathbb{C}$ defined as

$$\widetilde{f}(x,y) = \frac{1}{N^2}\sum_{\xi=0}^{N-1}\sum_{\eta=0}^{N-1} F(\xi,\eta)\mathrm{e}^{\frac{2\pi\mathrm{i}}{N}(x\xi+y\eta)}.$$

Corollary 3.7. Let $f(x,y)$ be a function $\{0,1,\dots,N-1\}^2 \to \mathbb{C}, N \in \mathbb{N}$. Then for every $(x,y),(\xi,\eta) \in \{0,1,\dots,N-1\}^2$ and $k,l \in \mathbb{Z}$ it holds:

$$f(x,y) = \widetilde{f}(x+kN, y+lN),$$

$$F(\xi,\eta) = \widetilde{F}(\xi+kN, \eta+lN).$$

In particular,

$$\widetilde{f}(x,y) = f(x,y), \qquad \widetilde{f}(-x,-y) = f(N-x, N-y),$$
$$\widetilde{F}(\xi,\eta) = F(\xi,\eta), \qquad \widetilde{F}(-\xi,-\eta) = F(N-\xi, N-\eta).$$

Proof. The claim is clear from formulae (3.1), (3.2) on page 37, the Fourier Inversion Theorem (Thm. 3.4) and the fact that goniometric functions are 2π-periodic. □

Definition 3.8 (Discrete Fourier transform of periodized functions). Let $f(x,y)$ be a function $\{0,1,\dots,N-1\}^2 \to \mathbb{C}, N \in \mathbb{N}$. The discrete Fourier transform of the periodization of function f, $\widetilde{f}(x,y) : \mathbb{Z}^2 \to \mathbb{C}$ is function $\mathcal{D}\{\widetilde{f}\}(\xi,\eta) = F(\xi,\eta) : \{0,1,\dots,N-1\}^2 \to \mathbb{C}$ defined as

$$F(\xi,\eta) = \sum_{x=0}^{N-1}\sum_{y=0}^{N-1} \widetilde{f}(x,y)\mathrm{e}^{-\frac{2\pi\mathrm{i}}{N}(x\xi+y\eta)}.$$

Definition 3.9 (Inverse discrete Fourier transform of periodized functions). Let $f(x,y)$ be a function $\{0,1,\dots,N-1\}^2 \to \mathbb{C}, N \in \mathbb{N}$ and let $F(\xi,\eta)$ be its discrete Fourier transform with periodization $\widetilde{F}(\xi,\eta) : \mathbb{Z}^2 \to \mathbb{C}$. The inverse discrete Fourier transform of function $\widetilde{F}(\xi,\eta)$ is function $\mathcal{D}^{-1}\{\widetilde{F}\}(x,y) : \{0,1,\dots,N-1\}^2 \to \mathbb{C}$ defined as

$$\mathcal{D}^{-1}\{\widetilde{F}\}(x,y) = \frac{1}{N^2}\sum_{\xi=0}^{N-1}\sum_{\eta=0}^{N-1} \widetilde{F}(\xi,\eta)\mathrm{e}^{\frac{2\pi\mathrm{i}}{N}(x\xi+y\eta)}.$$

Corollary 3.10. Let $f(x,y)$ be a function $\{0,1,\dots,N-1\}^2 \to \mathbb{C}, N \in \mathbb{N}$ with Fourier spectrum $F(\xi,\eta)$. For every $(x,y) \in \{0,1,\dots,N-1\}^2$, it holds:

$$\mathcal{D}\left\{f(x,y)\right\} = \mathcal{D}\left\{\widetilde{f}(x,y)\right\},$$

$$\mathcal{D}^{-1}\left\{\mathcal{D}\left\{\widetilde{f}(x,y)\right\}\right\} = \mathcal{D}^{-1}\left\{\widetilde{F}(\xi,\eta)\right\} = f(x,y)$$

Proof. The claim is a consequence of Definitions 3.8 and 3.9 and Theorem 3.4 (Fourier Inversion Theorem). □

Corollary 3.11. Let $f(x,y)$ be a function $\{0,1,\dots,N-1\}^2 \to \mathbb{C}, N \in \mathbb{N}$ with Fourier spectrum $F(\xi,\eta)$. For every $k,l \in \mathbb{Z}$ it holds:

$$\mathcal{D}\left\{f(x,y)\right\} = \sum_{x=k}^{k+N-1}\sum_{y=l}^{l+N-1} \widetilde{f}(x,y)\mathrm{e}^{-\frac{2\pi\mathrm{i}}{N}(x\xi+y\eta)},$$

$$\mathcal{D}^{-1}\left\{F(\xi,\eta)\right\} = \frac{1}{N^2}\sum_{\xi=k}^{k+N-1}\sum_{\eta=l}^{l+N-1} \widetilde{F}(\xi,\eta)\mathrm{e}^{\frac{2\pi\mathrm{i}}{N}(x\xi+y\eta)}.$$

Proof. The first claim is a consequence of the fact that both functions $\widetilde{f}$ and $\mathrm{e}^{-\frac{2\pi\mathrm{i}}{N}(x\xi+y\eta)}$ for fixed $\xi, \eta \in \mathbb{Z}$ are N-periodic. The second claim is a consequence of the fact that both functions $\widetilde{F}$ and $\mathrm{e}^{\frac{2\pi\mathrm{i}}{N}(x\xi+y\eta)}$ for fixed $x, y \in \mathbb{Z}$ are N-periodic. □

Definition 3.12 (Amplitude spectrum, phase spectrum). Let $f(x, y)$ be a function $\{0, 1, \ldots, N-1\}^2 \to \mathbb{C}, N \in \mathbb{N}$ with Fourier spectrum $F(\xi, \eta)$. The *amplitude spectrum* [8] of function f is function $A(\xi, \eta) : \{0, 1, \ldots, N-1\}^2 \to \mathbb{R}$ defined as

$$A(\xi, \eta) = |\mathcal{D}\{f(x, y)\}| = |F(\xi, \eta)|.$$

The *phase spectrum* [8] of function f is function $\Phi(\xi, \eta) : \{0, 1, \ldots, N-1\}^2 \to \langle 0, 2\pi)$ defined as

$$\begin{aligned} \mathfrak{Re}F(\xi, \eta) &= A(\xi, \eta)\cos\Phi(\xi, \eta), \\ \mathfrak{Im}F(\xi, \eta) &= A(\xi, \eta)\sin\Phi(\xi, \eta). \end{aligned}$$

If $A(\xi, \eta) = 0$ for some (ξ, η), we define $\Phi(\xi, \eta) = 0$.

In a similar way, we can define the amplitude and phase spectra of periodized functions and periodized amplitude and phase spectra.

3.2 Properties of the discrete Fourier transform

Table 3.1 shows the basic properties of the discrete Fourier transform. It is easy to prove them or they are proved below. Functions are listed on the left, their Fourier transforms on the right. $f(x, y), g(x, y)$ are functions $\{0, 1, \ldots, N-1\}^2 \to \mathbb{C}, N \in \mathbb{N}$, $\alpha, \beta, x_0, y_0, \xi_0, \eta_0$ are real constants, $f * g$ denotes discrete periodic convolution of functions f, g (described in Section 3.3). The formulae are taken from [8].

1.	$f(x, y)$	$F(\xi, \eta)$
2.	$\alpha f(x, y) + \beta g(x, y)$	$\alpha F(\xi, \eta) + \beta G(\xi, \eta)$
3.	$\widetilde{f}(x - x_0, y - y_0)$	$F(\xi, \eta) \cdot \mathrm{e}^{-\mathrm{i}(\xi x_0 + \eta y_0)}$
4.	$\mathrm{e}^{\mathrm{i}(\xi_0 x + \eta_0 y)} f(x, y)$	$\widetilde{F}(\xi - \xi_0, \eta - \eta_0)$
5.	$\widetilde{f}(-x, -y)$	$\widetilde{F}(-\xi, -\eta)$
6.	$f^*(x, y)$	$\widetilde{F^*}(-\xi, -\eta)$
7.	$(f * g)(x, y)$	$F(\xi, \eta) \cdot G(\xi, \eta)$
8.	$f(x, y) \cdot g(x, y)$	$\frac{1}{4\pi^2}(F * G)(\xi, \eta)$

Table 3.1: Basic properties of the discrete Fourier transform.

The second formula is linearity of the Fourier transform, the third formula is the Shift Theorem for Periodized Functions (Thm. 3.16), the fourth one is the Modulation Theorem [8], the fifth one is Theorem 3.14, the sixth one is Theorem 3.15 and the last two formulae are proved in Section 3.3.

Theorem 3.13. Let $f(x,y)$ be a function $\{0,1,\ldots,N-1\}^2 \to \mathbb{C}, N \in \mathbb{N}$ with Fourier spectrum $F(\xi,\eta)$. Then [15]

$$F(0,0) = \sum_{x=0}^{N-1}\sum_{y=0}^{N-1} f(x,y).$$

Proof.

$$F(0,0) = \sum_{x=0}^{N-1}\sum_{y=0}^{N-1} f(x,y)\mathrm{e}^{-\frac{2\pi\mathrm{i}}{N}(x\cdot 0+y\cdot 0)} = \sum_{x=0}^{N-1}\sum_{y=0}^{N-1} f(x,y)$$

□

Theorem 3.14. Let $f(x,y)$ be a function $\{0,1,\ldots,N-1\}^2 \to \mathbb{C}, N \in \mathbb{N}$ with Fourier spectrum $F(\xi,\eta)$. The Fourier transform of function f with reversed axes is function F with reversed axes, the inverse Fourier transform of function F with reversed axes is function f with reversed axes, i.e. [8]

$$\mathcal{D}\left\{\widetilde{f}(-x,-y)\right\} = \widetilde{F}(-\xi,-\eta) = F(N-\xi,N-\eta),$$

$$\mathcal{D}^{-1}\left\{\widetilde{F}(-\xi,-\eta)\right\} = \widetilde{f}(-x,-y) = f(N-x,N-y).$$

Proof.

$$\begin{aligned}\mathcal{D}\left\{\widetilde{f}(-x,-y)\right\} &= \sum_{x=0}^{N-1}\sum_{y=0}^{N-1}\widetilde{f}(-x,-y)\mathrm{e}^{-\frac{2\pi\mathrm{i}}{N}(x\xi+y\eta)} = \left|\begin{array}{l} s=-x \\ t=-y \end{array}\right| = \\ &= \sum_{s=-N+1}^{0}\sum_{t=-N+1}^{0}\widetilde{f}(s,t)\mathrm{e}^{-\frac{2\pi\mathrm{i}}{N}(-s\xi-t\eta)} = \widetilde{F}(-\xi,-\eta)\end{aligned}$$

The last equality is due to Corollary 3.11. The second claim is obtained from the first one by applying discrete inverse Fourier transform to its both sides. On the contrary with the Fourier transform discussed in Chapter 2, this step always works due to the Fourier Inversion Theorem (Thm. 3.4, page 37). □

Theorem 3.15. Let $f(x,y)$ be a function $\{0,1,\ldots,N-1\}^2 \to \mathbb{C}, N \in \mathbb{N}$ and let $F(\xi,\eta)$ be its Fourier spectrum. Then the Fourier transform of the complex conjugate of function f is $\widetilde{F}^*(-\xi,-\eta)$ [15], i.e.

$$\mathcal{D}\left\{f^*(x,y)\right\} = \widetilde{F}^*(-\xi,-\eta).$$

Proof.

$$\begin{aligned}\mathcal{D}\left\{f^*(x,y)\right\} &= \sum_{x=0}^{N-1}\sum_{y=0}^{N-1} f^*(x,y)\mathrm{e}^{-\frac{2\pi\mathrm{i}}{N}(x\xi+y\eta)} = \sum_{x=0}^{N-1}\sum_{y=0}^{N-1} f^*(x,y)\mathrm{e}^{\frac{2\pi\mathrm{i}}{N}(-x\xi-y\eta)} = \\ &= \left(\sum_{x=0}^{N-1}\sum_{y=0}^{N-1} f(x,y)\mathrm{e}^{-\frac{2\pi\mathrm{i}}{N}(-x\xi-y\eta)}\right)^* = \widetilde{F}^*(-\xi,-\eta).\end{aligned}$$

□

Theorem 3.16 (Shift Theorem for Periodized Functions). Let $f_1(x,y)$ be a function $\{0,1,\dots,N-1\}^2 \to \mathbb{C}, N \in \mathbb{N}$ and let $F_1(\xi,\eta)$ be its Fourier spectrum. Let us suppose function $f_2(x,y) : \{0,1,\dots,N-1\}^2 \to \mathbb{C}$ such that

$$f_2(x,y) = \widetilde{f_1}(x-x_0, y-y_0),$$

where $x_0, y_0 \in \mathbb{Z}$ are given numbers. Let $F_2(\xi,\eta)$ be the Fourier spectrum of function $f_2(x,y)$. Then it holds

$$F_2(\xi,\eta) = \mathrm{e}^{-\frac{2\pi\mathrm{i}}{N}(\xi x_0+\eta y_0)} F_1(\xi,\eta).$$

Proof. The following proof is a generalized version of the one-dimensional case proved in [15].

$$\begin{aligned} F_2(\xi,\eta) &= \sum_{x=0}^{N-1}\sum_{y=0}^{N-1} f_2(x,y)\mathrm{e}^{-\frac{2\pi\mathrm{i}}{N}(x\xi+y\eta)} = \sum_{x=0}^{N-1}\sum_{y=0}^{N-1} \widetilde{f_1}(x-x_0,y-y_0)\mathrm{e}^{-\frac{2\pi\mathrm{i}}{N}(x\xi+y\eta)} = \\ &= \left| \begin{array}{l} s = x - x_0 \\ t = y - y_0 \end{array} \right| = \sum_{s=-x_0}^{N-1-x_0}\sum_{t=-y_0}^{N-1-y_0} \widetilde{f_1}(s,t)\mathrm{e}^{-\frac{2\pi\mathrm{i}}{N}[\xi(s+x_0)+\eta(t+y_0)]} = \\ &= \mathrm{e}^{-\frac{2\pi\mathrm{i}}{N}(\xi x_0+\eta y_0)} \sum_{s=-x_0}^{N-1-x_0}\sum_{t=-y_0}^{N-1-y_0} \widetilde{f_1}(s,t)\mathrm{e}^{-\frac{2\pi\mathrm{i}}{N}(s\xi+t\eta)} = \mathrm{e}^{-\frac{2\pi\mathrm{i}}{N}(\xi x_0+\eta y_0)} F_1(\xi,\eta) \end{aligned}$$

The last equality is due to Corollary 3.11. □

Theorem 3.17 (Shift Theorem). Let $f_1(x,y)$ be a function $\{0,1,\dots,N-1\}^2 \to \mathbb{C}, N \in \mathbb{N}$. Let $x_0, y_0 \in \mathbb{Z},\ k,l \in \mathbb{N}_0, M \in \mathbb{N}$ are given numbers such that

$$M < N,$$

$$\begin{aligned} k + x_0 \geq 0, \qquad & k + M + x_0 \leq N-1, \\ l + y_0 \geq 0, \qquad & l + M + y_0 \leq N-1, \end{aligned}$$

and let f_1 be zero out of $\{k,k+1,\dots,k+M-1\} \times \{l,l+1,\dots,l+M-1\}$. Let $f_2(x,y)$ be a function $\{0,1,\dots,N-1\}^2 \to \mathbb{C}$ such that

$$f_2(x,y) = \begin{cases} f_1(x-x_0,y-y_0) & \text{if } \max\{0,x_0\} \leq x \leq \min\{N-1, N-1+x_0\}, \\ & \quad \max\{0,y_0\} \leq y \leq \min\{N-1, N-1+y_0\}, \\ 0 & \text{else.} \end{cases}$$

Such relations are illustrated in Figure 3.1. Let $F_1(\xi,\eta), F_2(\xi,\eta)$ be the Fourier spectra of functions f_1 and f_2. Then it holds

$$F_2(\xi,\eta) = \mathrm{e}^{-\frac{2\pi\mathrm{i}}{N}(\xi x_0+\eta y_0)} F_1(\xi,\eta).$$

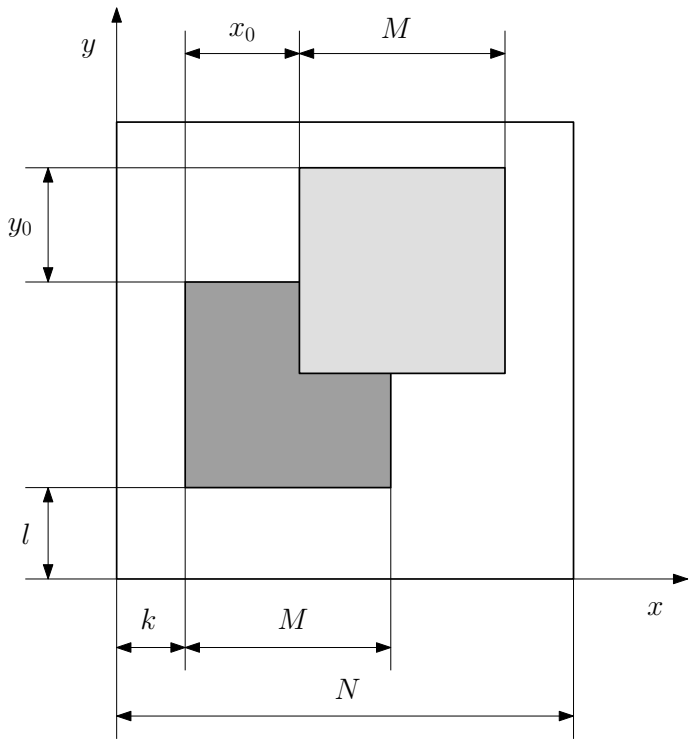


Figure 3.1: Illustration of relations between functions f_1 and f_2 and numbers k, l, M, N, x_0, y_0 in the Shift Theorem (Thm. 3.17). Function f_1 is zero outside the darker square, function f_2 outside the brighter square. x_0,y_0 are positive numbers in the picture.

Proof.

$$
\begin{aligned}
F_2(\xi,\eta) &= \sum_{x=0}^{N-1}\sum_{y=0}^{N-1} f_2(x,y)\mathrm{e}^{-\frac{2\pi\mathrm{i}}{N}(x\xi+y\eta)} = \sum_{x=k+x_0}^{k+M-1+x_0}\sum_{y=l+y_0}^{l+M-1+y_0} f_2(x,y)\mathrm{e}^{-\frac{2\pi\mathrm{i}}{N}(x\xi+y\eta)} = \\
&= \sum_{x=k+x_0}^{k+M-1+x_0}\sum_{y=l+y_0}^{l+M-1+y_0} f_1(x-x_0,y-y_0)\mathrm{e}^{-\frac{2\pi\mathrm{i}}{N}(x\xi+y\eta)} = \\
&= \left|\begin{array}{l} s = x - x_0 \\ t = y - y_0 \end{array}\right| = \sum_{s=k}^{k+M-1}\sum_{t=l}^{l+M-1} f_1(s,t)\mathrm{e}^{-\frac{2\pi\mathrm{i}}{N}[\xi(s+x_0)+\eta(t+y_0)]} = \\
&= \mathrm{e}^{-\frac{2\pi\mathrm{i}}{N}(\xi x_0+\eta y_0)}\sum_{s=k}^{k+M-1}\sum_{t=l}^{l+M-1} f_1(s,t)\mathrm{e}^{-\frac{2\pi\mathrm{i}}{N}(s\xi+t\eta)} = \mathrm{e}^{-\frac{2\pi\mathrm{i}}{N}(\xi x_0+\eta y_0)}F_1(\xi,\eta)
\end{aligned}
$$

□

Theorem 3.18. Let $f(x,y)$ be a function $\{0,1,\ldots,N-1\}^2 \to \mathbb{C}, N \in \mathbb{N}$. Then it holds

(a) $$\mathcal{D}\left\{\mathcal{D}\left\{f(x,y)\right\}\right\} = N^2\widetilde{f}(-x,-y),$$

(b) $$\mathcal{D}^{-1}\left\{\mathcal{D}^{-1}\left\{f(x,y)\right\}\right\} = \frac{1}{N^2}\widetilde{f}(-x,-y),$$

(c) $$\mathcal{D}\left\{f(x,y)\right\} = N^2\mathcal{D}^{-1}\left\{\widetilde{f}(-x,-y)\right\},$$

(d) $$\mathcal{D}^{-1}\left\{f(x,y)\right\} = \frac{1}{N^2}\mathcal{D}\left\{\widetilde{f}(-x,-y)\right\}$$

Proof.

(a)

$$\begin{aligned}\mathcal{D}\left\{\mathcal{D}\left\{f(x,y)\right\}\right\} &= \sum_{\xi=0}^{N-1}\sum_{\eta=0}^{N-1} F(\xi,\eta)\mathrm{e}^{-\frac{2\pi\mathrm{i}}{N}(x\xi+y\eta)} = \sum_{\xi=0}^{N-1}\sum_{\eta=0}^{N-1} F(\xi,\eta)\mathrm{e}^{\frac{2\pi\mathrm{i}}{N}(-x\xi-y\eta)} = \\ &= \left|\begin{array}{l}\sigma=-\xi\\ \tau=-\eta\end{array}\right| = \sum_{\sigma=-N+1}^{0}\sum_{\tau=-N+1}^{0} \widetilde{F}(-\sigma,-\tau)\mathrm{e}^{\frac{2\pi\mathrm{i}}{N}(x\sigma+y\tau)} = \\ &= N^2\mathcal{D}^{-1}\left\{\mathcal{D}\left\{\widetilde{f}(-x,-y)\right\}\right\} = N^2\widetilde{f}(-x,-y)\end{aligned}$$

(b)

$$\begin{aligned}\mathcal{D}^{-1}\left\{\mathcal{D}^{-1}\left\{f(x,y)\right\}\right\} &= \frac{1}{N^2}\sum_{\xi=0}^{N-1}\sum_{\eta=0}^{N-1}\left(\frac{1}{N^2}\sum_{s=0}^{N-1}\sum_{t=0}^{N-1} f(s,t)\mathrm{e}^{\frac{2\pi\mathrm{i}}{N}(s\xi+t\eta)}\right)\mathrm{e}^{\frac{2\pi\mathrm{i}}{N}(x\xi+y\eta)} = \\ &= \frac{1}{N^4}\sum_{\xi=0}^{N-1}\sum_{\eta=0}^{N-1}\left(\sum_{s=0}^{N-1}\sum_{t=0}^{N-1} f(s,t)\mathrm{e}^{-\frac{2\pi\mathrm{i}}{N}(-s\xi-t\eta)}\right)\mathrm{e}^{\frac{2\pi\mathrm{i}}{N}(x\xi+y\eta)} = \\ = \left|\begin{array}{l}u=-s\\ v=-t\end{array}\right| &= \frac{1}{N^4}\sum_{\xi=0}^{N-1}\sum_{\eta=0}^{N-1}\left(\sum_{u=-N+1}^{0}\sum_{v=-N+1}^{0}\widetilde{f}(-u,-v)\mathrm{e}^{-\frac{2\pi\mathrm{i}}{N}(u\xi+v\eta)}\right)\mathrm{e}^{\frac{2\pi\mathrm{i}}{N}(x\xi+y\eta)} = \\ &= \frac{1}{N^2}\mathcal{D}^{-1}\left\{\mathcal{D}\left\{\widetilde{f}(-x,-y)\right\}\right\} = \frac{1}{N^2}\widetilde{f}(-x,-y)\end{aligned}$$

(c) is obtained from (a) by applying the inverse Fourier transform to both its sides.

(d) is obtained from (c) by substituting $-x$ for x, $-y$ for y and dividing both sides by N^2.

□

Theorem 3.19. Let $f(x,y)$ be a function $\{0,1,\ldots,N-1\}^2 \to \mathbb{C}, N \in \mathbb{N}$ and have Fourier spectrum $F(\xi,\eta)$. f is a real function, i.e. $f(x,y) = f^*(x,y)\ \forall(x,y) \in \{0,1,\ldots,N-1\}^2$, if and only if $F(\xi,\eta) = \widetilde{F}^*(-\xi,-\eta)$ (partially stated in [15]).

Proof.

(a) Let us suppose that f is a real function. Then Theorem 3.15 (page 41) implies that

$$F(\xi,\eta) = \mathcal{D}\left\{f(x,y)\right\} = \mathcal{D}\left\{f^*(x,y)\right\} = F^*(-\xi,-\eta).$$

(b) Let us suppose that $F(\xi,\eta) = F^*(-\xi,-\eta)$. Then Theorem 3.15 and the Fourier Inversion Theorem (Thm. 3.4, page 37) imply that

$$f(x,y) = \mathcal{D}^{-1}\left\{F(\xi,\eta)\right\} = \mathcal{D}^{-1}\left\{\widetilde{F}^*(-\xi,-\eta)\right\} = \mathcal{D}^{-1}\left\{\mathcal{D}\left\{f^*(x,y)\right\}\right\} = f^*(x,y)$$

□

Corollary 3.20. Let function $f(x,y) : \{0,1,\ldots,N-1\}^2 \to \mathbb{R}, N \in \mathbb{N}$. Then

$$A(\xi,\eta) = A(-\xi,-\eta).$$

Proof. Consequence of Theorem 3.19. □

Corollary 3.21. Let function $f(x,y) : \{0,1,\ldots,N-1\}^2 \to \mathbb{R}, N \in \mathbb{N}$. Let function $G(\xi,\eta) : \{0,1,\ldots,N-1\}^2 \to \mathbb{R}$ such that $G(\xi,\eta) = \widetilde{G}(-\xi,-\eta)$. Then

$$\mathcal{D}^{-1}\left\{F(\xi,\eta)\cdot G(\xi,\eta)\right\}$$

is real.

Proof. According to Theorem 3.19, if f is real, then

$$F(\xi,\eta) = \widetilde{F}^*(-\xi,-\eta).$$

Multiplying the equality by G, we obtain

$$F(\xi,\eta)\cdot G(\xi,\eta) = \widetilde{F}^*(-\xi,-\eta)\cdot\widetilde{G}(-\xi,-\eta) = (\widetilde{F}(-\xi,-\eta)\cdot\widetilde{G}(-\xi,-\eta))^*.$$

Then again according to Theorem 3.19,

$$\mathcal{D}^{-1}\left\{F(\xi,\eta)\cdot G(\xi,\eta)\right\}$$

is real. □

3.3 Discrete periodic convolution

Definition 3.22 (Discrete periodic convolution). Let $f_1(x,y), f_2(x,y)$ be functions $\{0,1,\ldots,N-1\}^2 \to \mathbb{C}, N \in \mathbb{N}$. Function $f(x,y) : \{0,1,\ldots,N-1\}^2 \to \mathbb{C}$ is called the *discrete periodic convolution* [8] of functions f_1, f_2, denoted by $f(x,y) = f_1(x,y) * f_2(x,y)$, if

$$f(x,y) = \sum_{s=0}^{N-1}\sum_{t=0}^{N-1} f_1(s,t)\widetilde{f_2}(x-s,y-t).$$

Theorem 3.23. Let functions $f_1(x,y), f_2(x,y) : \{0,1,\ldots,N-1\}^2 \to \mathbb{C}, N \in \mathbb{N}$ have Fourier spectra $F_1(\xi,\eta), F_2(\xi,\eta)$. Then

$$\mathcal{D}\left\{f_1(x,y) * f_2(x,y)\right\} = F_1(\xi,\eta)\cdot F_2(\xi,\eta).$$

Proof. A similar proof can be found in [8]. Let $f(x,y) = f_1(x,y) * f_2(x,y)$. Then

$$\begin{aligned}
\mathcal{D}\{f(x,y)\} &= \sum_{x=0}^{N-1}\sum_{y=0}^{N-1}\left(\sum_{s=0}^{N-1}\sum_{t=0}^{N-1} f_1(s,t)\widetilde{f_2}(x-s,y-t)\right)\mathrm{e}^{-\frac{2\pi\mathrm{i}}{N}(x\xi+y\eta)} = \\
&= \sum_{s=0}^{N-1}\sum_{t=0}^{N-1} f_1(s,t)\sum_{x=0}^{N-1}\sum_{y=0}^{N-1}\widetilde{f_2}(x-s,y-t)\mathrm{e}^{-\frac{2\pi\mathrm{i}}{N}(x\xi+y\eta)} = \left|\begin{array}{l} p = x-s \\ q = y-t \end{array}\right| = \\
&= \sum_{s=0}^{N-1}\sum_{t=0}^{N-1} f_1(s,t)\sum_{p=-s}^{N-1-s}\sum_{q=-t}^{N-1-t}\widetilde{f_2}(p,q)\mathrm{e}^{-\frac{2\pi\mathrm{i}}{N}[\xi(p+s)+\eta(q+t)]} = \\
&= \sum_{s=0}^{N-1}\sum_{t=0}^{N-1} f_1(s,t)\mathrm{e}^{-\frac{2\pi\mathrm{i}}{N}(s\xi+t\eta)}\cdot\sum_{p=-s}^{N-1-s}\sum_{q=-t}^{N-1-t}\widetilde{f_2}(p,q)\mathrm{e}^{-\frac{2\pi\mathrm{i}}{N}(p\xi+q\eta)} = \\
&= F_1(\xi,\eta)\cdot F_2(\xi,\eta).
\end{aligned}$$

□

Theorem 3.24. Let functions $f_1(x,y), f_2(x,y) : \{0,1,\ldots,N-1\}^2 \to \mathbb{C}, N \in \mathbb{N}$ have Fourier spectra $F_1(\xi,\eta), F_2(\xi,\eta)$. Then [15]

$$\mathcal{D}\{f_1(x,y)\cdot f_2(x,y)\} = \frac{1}{N^2}F_1(\xi,\eta) * F_2(\xi,\eta).$$

Proof. Let $f(x,y) = f_1(x,y) * f_2(x,y)$. Using the Fourier Inversion Theorem 3.4 (page 37) we can compute $\mathcal{D}\{f(x,y)\}$ as

$$\begin{aligned}
\mathcal{D}\{f(x,y)\} &= \sum_{x=0}^{N-1}\sum_{y=0}^{N-1} f_1(x,y)f_2(x,y)\mathrm{e}^{-\frac{2\pi\mathrm{i}}{N}(x\xi+y\eta)} = \\
&= \sum_{x=0}^{N-1}\sum_{y=0}^{N-1}\left(\frac{1}{N^2}\sum_{\sigma=0}^{N-1}\sum_{\tau=0}^{N-1} F_1(\sigma,\tau)\mathrm{e}^{\frac{2\pi\mathrm{i}}{N}(x\sigma+y\tau)}\right) f_2(x,y)\mathrm{e}^{-\frac{2\pi\mathrm{i}}{N}(x\xi+y\eta)} = \\
&= \frac{1}{N^2}\sum_{\sigma=0}^{N-1}\sum_{\tau=0}^{N-1} F_1(\sigma,\tau)\sum_{x=0}^{N-1}\sum_{y=0}^{N-1} f_2(x,y)\mathrm{e}^{-\frac{2\pi\mathrm{i}}{N}[x(\xi-\sigma)+y(\eta-\tau)]} = \\
&= \frac{1}{N^2}\sum_{\sigma=0}^{N-1}\sum_{\tau=0}^{N-1} F_1(\sigma,\tau)\widetilde{F_2}(\xi-\sigma,\eta-\tau) = \frac{1}{N^2}F_1(\xi,\eta) * F_2(\xi,\eta).
\end{aligned}$$

□

3.4 Cross correlation, phase correlation

Definition 3.25 (Cross-power spectrum, normalized cross-power spectrum, semi-normalized cross-power spectrum). Let functions $f_1(x,y), f_2(x,y) : \{0,1,\ldots, N-1\}^2 \to \mathbb{C}, N \in \mathbb{N}$ and Fourier spectra $F_1(\xi,\eta), F_2(\xi,\eta)$. The *cross-power spectrum* [23] of functions f_1, f_2 is function $C_{f_1,f_2}(\xi,\eta) : \{0,1,\ldots,N-1\}^2 \to \mathbb{C}$ defined by

$$C_{f_1,f_2}(\xi,\eta) = F_1(\xi,\eta)\cdot F_2^*(\xi,\eta).$$

The *normalized cross-power spectrum* [23] of functions f_1, f_2 is function $Z_{f_1,f_2}(\xi,\eta) : \{0,1,\ldots,N-1\}^2 \to \mathbb{C}$ defined by

$$Z_{f_1,f_2}(\xi,\eta) = \frac{F_1(\xi,\eta)\cdot F_2^*(\xi,\eta)}{|F_1(\xi,\eta)\cdot F_2(\xi,\eta)|}.$$

If $p,q \in \mathbb{R}^+$ are given constants, the *semi-normalized cross-power spectrum* of functions f_1, f_2 with parameters p, q is function $Z^{p,q}_{f_1,f_2}(\xi,\eta) : \{0,1,\ldots,N-1\}^2 \to \mathbb{C}$ defined as

$$Z^{p,q}_{f_1,f_2}(\xi,\eta) = \frac{F_1(\xi,\eta)\cdot F_2^*(\xi,\eta)}{(|F_1(\xi,\eta)|+p)\cdot(|F_2(\xi,\eta)|+q)}.$$

Definition 3.26 (Cross-correlation function, phase-correlation function, semi-phase correlation function). Let functions $f_1(x,y), f_2(x,y) : \{0,1,\ldots,N-1\}^2 \to \mathbb{C}, N \in \mathbb{N}$ have Fourier spectra $F_1(\xi,\eta)$, $F_2(\xi,\eta)$. Function $Q_{f_1,f_2}(x,y) : \{0,1,\ldots,N-1\}^2 \to \mathbb{C}$ defined as

$$Q_{f_1,f_2}(x,y) = \mathcal{D}^{-1}\left\{C_{f_1,f_2}(\xi,\eta)\right\} = \mathcal{D}^{-1}\left\{F_1(\xi,\eta)\cdot F_2^*(\xi,\eta)\right\}$$

is called the *cross-correlation function* of functions f_1, f_2. Function $P_{f_1,f_2}(x,y) : \{0,1,\ldots,N-1\}^2 \to \mathbb{C}$ defined as

$$P_{f_1,f_2}(x,y) = \mathcal{D}^{-1}\left\{Z_{f_1,f_2}(\xi,\eta)\right\} = \mathcal{D}^{-1}\left\{\frac{F_1(\xi,\eta)\cdot F_2^*(\xi,\eta)}{|F_1(\xi,\eta)\cdot F_2(\xi,\eta)|}\right\}$$

is called the *phase-correlation function* of functions f_1, f_2 and function $P^{p,q}_{f_1,f_2}(x,y) : \{0,1,\ldots,N-1\}^2 \to \mathbb{C}$ defined as

$$P^{p,q}_{f_1,f_2}(x,y) = \mathcal{D}^{-1}\left\{Z^{p,q}_{f_1,f_2}(\xi,\eta)\right\} = \mathcal{D}^{-1}\left\{\frac{F_1(\xi,\eta)\cdot F_2^*(\xi,\eta)}{(|F_1(\xi,\eta)|+p)\cdot(|F_2(\xi,\eta)|+q)}\right\}$$

is called the *semi-phase correlation function* of functions f_1, f_2 with parameters p, q.

Theorem 3.27 (Cross-correlation function for real functions). Let $f_1(x,y), f_2(x,y)$ be functions $\{0,1,\ldots,N-1\}^2 \to \mathbb{R}, N \in \mathbb{N}$ (real functions) and have Fourier spectra $F_1(\xi,\eta)$, $F_2(\xi,\eta)$. Then the cross-correlation function of these functions is real.

Proof. Using Theorems 3.4 (page 37) and 3.15 (page 41) we can compute the cross correlation function as

$$\begin{aligned} Q_{f_1,f_2}(x,y) &= \mathcal{D}^{-1}\left\{F_1(\xi,\eta)\cdot F_2^*(\xi,\eta)\right\} = \mathcal{D}^{-1}\left\{\mathcal{D}\left\{f_1(x,y)\right\}\cdot\mathcal{D}\left\{f_2^*(-x,-y)\right\}\right\} = \\ &= \mathcal{D}^{-1}\left\{\mathcal{D}\left\{f_1(x,y) * f_2^*(-x,-y)\right\}\right\} = f_1(x,y) * f_2^*(-x,-y) = \\ &= f_1(x,y) * f_2(-x,-y), \end{aligned}$$

which is a real function. □

Remark 3.28. From [15], generalized into two dimensions and with notation from Remark 2.65 (page 35), the *discrete periodical correlation* is a function

$$f_1(x,y) \star f_2(x,y) = \sum_{s=0}^{N-1}\sum_{t=0}^{N-1} f_1^*(s,t)\widetilde{f_2}(s+x,t+y).$$

Using [15], we can prove a formula to transform cross-correlation to convolution

$$f_1(x,y) \star f_2(x,y) = \widetilde{f_1^*}(-x,-y) * f_2(x,y).$$

Therefore, using the proof of Theorem 3.27 and commutativity of convolution [15]

$$Q_{f_1,f_2}(x,y) = f_1(x,y) * \widetilde{f_2^*}(-x,-y) = \widetilde{f_2^*}(-x,-y) * f_1(x,y) = f_2(x,y) \star f_1(x,y),$$

which gives the connection between the cross-correlation used in Fourier analysis [15] and the cross-correlation function Q_{f_1,f_2} defined in Definition 3.26 and used in image registration.

Theorem 3.29. Let functions $f_1(x,y),\ f_2(x,y) : \{0,1,\ldots,N-1\}^2 \to \mathbb{R}, N \in \mathbb{N}$ have amplitude spectra $A_1(\xi,\eta),\ A_2(\xi,\eta)$. Let $p,q \in \mathbb{R}_0^+$. If

$$A_1(\xi,\eta), A_2(\xi,\eta) > 0\ \forall(\xi,\eta) \quad \text{or} \quad p,q > 0$$

then the semi-phase correlation function of functions f_1, f_2 with parameters p,q

$$\mathcal{D}^{-1}\left\{\frac{F_1(\xi,\eta)\cdot F_2^*(\xi,\eta)}{(A_1(\xi,\eta)+p)\cdot(A_2(\xi,\eta)+q)}\right\}$$

is a real function. Especially, if $A_1(\xi,\eta), A_2(\xi,\eta) > 0\ \forall(\xi,\eta)$, the phase-correlation function of functions f_1, f_2 is real.

Proof. Amplitude spectra are real functions and from Corollary 3.20 (page 45) we have that

$$A_k(\xi,\eta) = \widetilde{A}_k(-\xi,-\eta) \ \text{ for } k = 1,2.$$

Furthermore, if

$$A(\xi,\eta) = (A_1(\xi,\eta)+p)\cdot(A_2(\xi,\eta)+q),$$

we have $A(\xi,\eta) = \widetilde{A}(-\xi,-\eta)$ and A is again real. Now function $\frac{1}{A(\xi,\eta)}$ has all properties of function G in Corollary 3.21 on page 45. It is a real bounded function with $A(\xi,\eta) = \widetilde{A}(-\xi,-\eta)$. Function $F_1(\xi,\eta)\cdot F_2^*(\xi,\eta)$ has the properties of function F in Corollary 3.21 due to the fact that the cross-correlation function of two real functions is real (Thm. 3.27). Therefore,

$$\mathcal{D}^{-1}\left\{\frac{F_1(\xi,\eta)\cdot F_2^*(\xi,\eta)}{A(\xi,\eta)}\right\} = \mathcal{D}^{-1}\left\{\frac{F_1(\xi,\eta)\cdot F_2^*(\xi,\eta)}{(A_1(\xi,\eta)+p)\cdot(A_2(\xi,\eta)+q)}\right\}$$

is real. If $A_1(\xi,\eta), A_2(\xi,\eta) > 0\ \forall(\xi,\eta)$, we can set $p = q = 0$ and conclude that the phase-correlation function of real functions is real. □

Definition 3.30 (Discrete impulse function). Let function $d(x,y)$ be defined on $\{0,1,\ldots,N-1\}^2$ as

$$d(x,y) = \begin{cases} 1 & \text{if } (x,y) = (0,0) \\ 0 & \text{else.} \end{cases}$$

Function d is called the *discrete impulse function*.

Theorem 3.31 (Phase-correlation function of shifted functions). Let functions f_1, f_2 fulfill the same assumptions as in the Shift Theorem (Thm. 3.17, page 42). Then the phase-correlation function of functions f_1, f_2 is the discrete impulse function d shifted by $(-x_0, -y_0)$

$$P_{f_1,f_2}(x,y) = \widetilde{d}(x+x_0, y+y_0).$$

Proof. Shift Theorem (Thm. 3.17) implies that

$$Z_{f_1,f_2}(\xi,\eta) = \frac{F_1(\xi,\eta)\cdot F_1^*(\xi,\eta)\left(\mathrm{e}^{-\frac{2\pi\mathrm{i}}{N}(\xi x_0+\eta y_0)}\right)^*}{\left|F_1(\xi,\eta)\cdot F_1(\xi,\eta)\mathrm{e}^{-\frac{2\pi\mathrm{i}}{N}(\xi x_0+\eta y_0)}\right|} = \mathrm{e}^{\frac{2\pi\mathrm{i}}{N}(\xi x_0+\eta y_0)}.$$

According to formula (3.18d) (since functions f_1 and F_1 are both $\{0,1,\ldots,N-1\}^2 \to \mathbb{C}$, we can interchange their roles and use formula (3.18d) for F_1),

$$\begin{aligned}
\mathcal{D}^{-1}\left\{Z_{f_1,f_2}(\xi,\eta)\right\} &= \frac{1}{N^2}\mathcal{D}\left\{\widetilde{Z}_{f_1,f_2}(-\xi,-\eta)\right\} = \frac{1}{N^2}\sum_{\xi=0}^{N-1}\sum_{\eta=0}^{N-1}\mathrm{e}^{-\frac{2\pi\mathrm{i}}{N}(\xi x_0+\eta y_0)}\mathrm{e}^{-\frac{2\pi\mathrm{i}}{N}(\xi x+\eta y)} = \\
&= \frac{1}{N^2}\sum_{\xi=0}^{N-1}\mathrm{e}^{-\frac{2\pi\mathrm{i}}{N}\xi(x+x_0)}\cdot\sum_{\eta=0}^{N-1}\mathrm{e}^{-\frac{2\pi\mathrm{i}}{N}\eta(y_0+y)} = \\
&= \frac{1}{N^2}\underbrace{\sum_{\xi=0}^{N-1}\left(\mathrm{e}^{-\frac{2\pi\mathrm{i}}{N}(x+x_0)}\right)^{\xi}}_{\text{denote by } g(x)}\cdot\underbrace{\sum_{\eta=0}^{N-1}\left(\mathrm{e}^{-\frac{2\pi\mathrm{i}}{N}(y_0+y)}\right)^{\eta}}_{\text{denote by } g(y).}
\end{aligned}$$

Similarly to the proof of the Fourier Inversion Theorem (Thm. 3.4, page 37), $g(x)$ is a finite geometrical series. For $x = -x_0 + kN$, where k an arbitrary integer, $g(x) = N$ since all elements of the series are equal to one. Otherwise,

$$g(x) = \frac{1-\left(\mathrm{e}^{-\frac{2\pi\mathrm{i}}{N}(x+x_0)}\right)^N}{1-\mathrm{e}^{-\frac{2\pi\mathrm{i}}{N}(x+x_0)}} = \frac{1-\mathrm{e}^{-2\pi\mathrm{i}(x+x_0)}}{1-\mathrm{e}^{-\frac{2\pi\mathrm{i}}{N}(x+x_0)}} = \frac{1-1}{1-\mathrm{e}^{-\frac{2\pi\mathrm{i}}{N}(x+x_0)}} = 0.$$

Analogically,

$$g(y) = \begin{cases} N & \text{if } y = -y_0 + lN, l \in \mathbb{Z} \\ 0 & \text{else.} \end{cases}$$

Hence,

$$\begin{aligned}
P_{f_1,f_2}(x,y) &= \mathcal{D}^{-1}\left\{Z_{f_1,f_2}(\xi,\eta)\right\} = \begin{cases} 1 & \text{if } (x,y) = (-x_0+kN, -y_0+lN) \\ & \text{for some } k,l \in \mathbb{Z} \\ 0 & \text{else} \end{cases} \\
&= \widetilde{d}(x+x_0, y+y_0).
\end{aligned}$$

□

Chapter 4

Image registration

4.1 Digital image

The chips of digital cameras have usually rectangular matrices of photo-sensitive elements. CMOS and CCD chips have photo-diodes as photo-sensitive elements. If there are color filters before the photo-diodes (organized in a so called Bayer mask [14]), some photo-diodes receive only a part of the spectrum and each pixel has information about one color channel only. The other color components are interpolated in each pixel to obtain a color image. There is another system of sensors for color digital photography, Foveon [28], where each photo-sensitive element captures information about all three color channels, the red, green, and the blue one.

It rarely makes sense to register color components separately. The only case when it makes sense is correction of chromatic abberation of the used optical system. Therefore, for image registration we convert digital color images into gray scale.

Definition 4.1 (Digital gray-scale image). Let $R = \{0, 1, \ldots, M-1\} \times \{0, 1, \ldots, N-1\}, M, N \in \mathbb{N}$ and let $W = \{0, 1, \ldots, w-1\}, w \in \mathbb{N}$. Function [20]

$$f(x, y) : R \to W$$

is called a *digital gray-scale image* or *image* only if no confusion may be caused. M is called the *image width*, N the *image height*. Elements of R are called *pixels*, value of f in pixel (x, y) is called the *pixel value*. The value of w determines the image *dynamic range*. We say that the *dynamic range is n bits per pixel* (it is an *n-bit image*) if $w = 2^n$.

Definition 4.2 (Digital color image). A *digital color image* is a triple of digital gray-scale images (r, g, b), which are called (in this order) the *red*, *green* and *blue color channel*.

Remark 4.3. Usually, matrices are used to represent images, they are called *image matrices*. These matrices are only tables of pixel values and operations defined for matrices do not make sense for them.

Definition 4.4 (Additive noise). Let f be a digital gray-scale image representing an ideal image (containing no additive noise). Let n be a digital gray-scale image of the same size as f whose pixel values are rounded independent realization of a random variable X. Let $h = f + n$. Then we say that image h *contains additive noise* [20]. Image n is called the *noise image* and the characteristics of X are called the *characteristics of additive noise*.

Additive noise usually has normal distribution. Its standard deviation is dependent on temperature of the sensor. The higher is the temperature the higher is the standard deviation of the impulse noise.

Remark 4.5. The previous definition does not say what happens to pixels where $f+n > w-1$ or $f+n<0$ ($w-1$ is the maximal possible pixel value, see Def. 4.1). Usually, we try to set the exposure on the camera so that such cases do not happen. If they do, all values higher than $w-1$ are mapped to $w-1$ and values lower than 0 are mapped to 0

$$h(x,y)=\begin{cases}0 & \text{if } f(x,y)+n(x,y)<0\\ f(x,y)+n(x,y) & \text{if } 0\leq f(x,y)+n(x,y)<w\\ w-1 & \text{if } f(x,y)+n(x,y)\geq w.\end{cases}$$

To convert a color image into a gray-scale one, we compute a convex combination of the red, green, and blue color channels

$$f(x,y)=\mathrm{Round}(c_r r(x,y)+c_g g(x,y)+c_b b(x,y)),$$

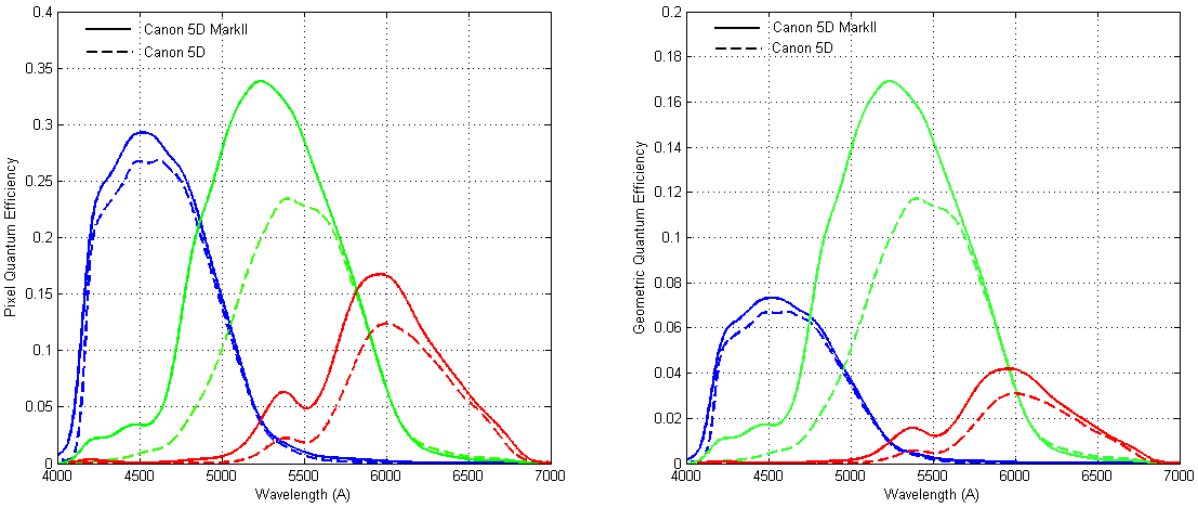
where $c_r, c_g, c_b \in \langle 0,1\rangle$ and $c_r+c_g+c_b=1$. The constants c_r, c_g, c_b should be chosen to minimize the standard deviation of additive noise in image f. There is no rule for choosing the constants that works for all images. For instance, if the scene of the image is mostly red (images taken with red illumination or through a narrow-band filter), c_r should be much higher than c_g and c_b. For general images, the constants should be chosen with respect to the number of each color-sensitive sensors and their pixel quantum efficiency, thus the geometric quantum efficiency of the sensor (see Fig. 4.1). In the best case, the constants should be chosen as multiples of the integral of quantum efficiency over the whole spectrum. However, for a given camera, it is very hard to find its quantum efficiency. The producers do not publish them and results of measurements are available for a few cameras only. Then, we must use some assessments, e.g.

$$c_r=\frac{1}{9},\quad c_g=\frac{6}{9},\quad c_b=\frac{2}{9}.$$

If we want to register astronomical images, which are taken through thicker layers of the atmosphere, it is better to increase c_r and decrease c_b, because the atmosphere absorbs blue light more than red light.

Remark 4.6 (Connection between $\mathcal{F}$ and $\mathcal{D}$). [8] very well describes the connection between the Fourier transform of periodic functions or functions defined on a cartesian product of two finite intervals, its Fourier series and discrete Fourier transform. This enables us to work with gray-scale images and functions applied on them (by multiplication, convolution) as with functions defined on $\mathbb{R}^2$. This approach makes many descriptions possible to be written in a compact form and some computations to be performed.

Remark 4.7. Some theorems about the properties of the Fourier transform in Chapter 2 require that function f is continuous and its Fourier spectrum F is in $\mathcal{L}(\mathbb{R}^2)$. If we consider a digital image as a function $\mathbb{R}^2\to\mathbb{C}$ (used in symbolic computations especially with rotated and scaled images), we know its values in a finite number of points only. The functions can be made as smooth as we want them to be and since the function values are defined on a bounded set only, the functions can also be taken as being in $\mathcal{L}\,(\mathbb{R}^2)$.



(a) (b)

Figure 4.1: Quantum efficiency of digital cameras Canon EOS 5D and 5D Mark II taken from [50]. The pixel quantum efficiency (a) is the efficiency of an isolated pixel, which may be either red-sensitive, green-sensitive, or blue-sensitive drawn with the respective color. Geometric quantum efficiency (b) takes into account the organization of the Bayer mask of the detector. It shows the efficiency of a 2×2 pixel elementary structure. There is one pixel for the red channel, two pixels for the green channel and one pixel for the blue channel. Therefore, the geometric quantum efficiency is reduced by factors of 0.25, 0.50 and 0.25 with respect to pixel quantum efficiency.

4.2 Registration of idealized images

In this section, we will focus on registration of images which are identical up to a shift, scale-change or rotation. They neither contain different noise nor have different brightness or contrast. The images which the methods will be demonstrated on will be cropped from a bigger image so that they are identical up to a shift, scale-change or rotation.

4.2.1 Registration of identical images

Firstly, if we consider registration of functions in $\mathcal{L}(\mathbb{R}^2)$ discussed in Section 2.6 (from page 34), the phase-correlation function of two identical functions is the Dirac distribution. However, the normalized cross-power spectrum is only defined for such values of ξ and η where the Fourier spectrum of the image to be registered is non-zero (Def. 2.62, page 34). Otherwise, the fraction defining the normalized cross-power spectrum is undefined. For the integral used in the inverse Fourier transform defining the phase-correlation function (Def. 2.63, page 34), it does not matter if the integrand value is undefined in a finite number of points, but it is typically not the case. If the cross-power spectrum is zero in a point (ξ, η), it is physically reasonable that it is also zero in a neighborhood of the point. Most typically, the Fourier spectrum (and therefore also the cross-power spectrum) has a bounded support which is located around $(0, 0)$.

In the practical case, registration of digital images, the phase-correlation function of

two identical images is the discrete impulse function (Def. 3.30, page 48). The case that the fraction defining the normalized cross-power spectrum is zero (Def. 3.25, page 46) is fatal for the sum used in the inverse Fourier transform defining the phase-correlation function (Def. 3.26, page 47). It happens if the registered image does not have the highest spatial frequencies, the Fourier spectrum of such image is zero in points (ξ, η) far from the origin. For instance, it is an ideal image containing no additive noise and the image has "empty resolution", it is an artificially enlarged replica of an image. In reality, every image taken by a digital camera contains additive noise, which ensures that no values of the spectrum are exactly zero. However, they may be very low causing very high values of $Z_{f,f}$.

On the other hand, the cross-correlation function of two identical functions, is a function which does not set any requirements on the functions. It has a global maximum in $(0, 0)$ [1] and its shape depends on the function, since Theorem 3.27 on page 47 says that

$$Q_{f,f}(x, y) = f(x, y) * f(-x, -y).$$

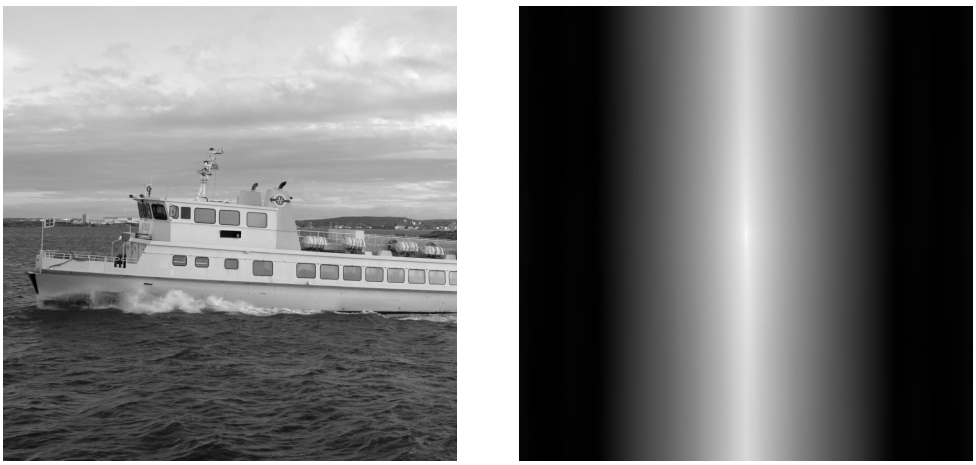
Figure 4.2 shows that the cross-correlation function even need not be concave and thus may have other local maxima than $(0, 0)$.

The problem of division by zero or "nearly zero" in phase-correlation can be solved by using a semi-phase correlation function (Definitions 2.62 on page 34 and 3.25 on page 46). The selection of numbers p and q will be discussed in Section 4.3.1 (starting from page 65).

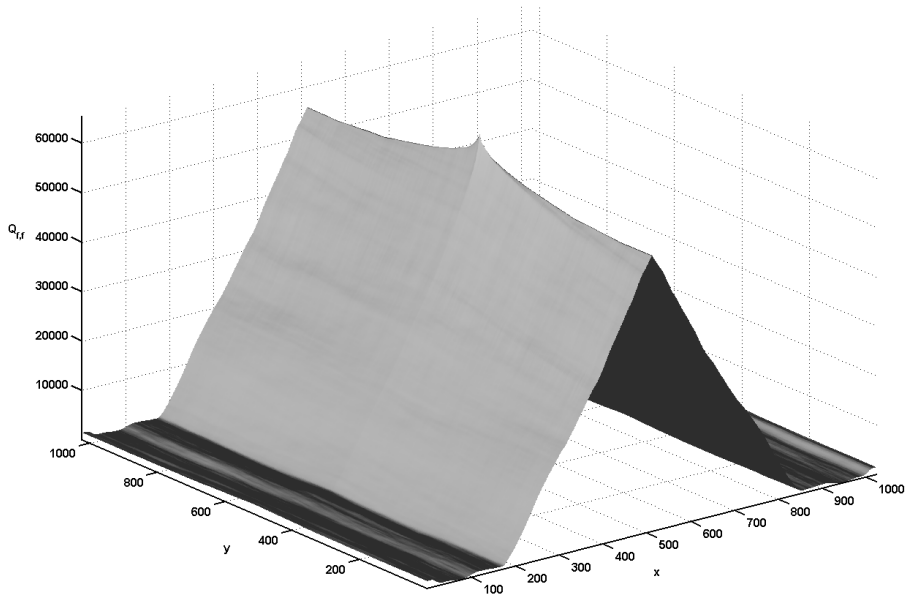
Remark 4.8. All figures showing a Fourier spectrum or a cross-correlation function or a phase-correlation function in this thesis are so called permuted. The images are considered as periodized in the sense of Definition 3.6 and then the figures show the values of the periodization from $-\frac{N}{2}$ to $\frac{N}{2} - 1$. The reason for the permutation is that point $(0, 0)$ is drawn in the middle and the graph of the cross-correlation function or the phase-correlation function resembles more the graph of a function defined on $\mathbb{R}^2$. Here we assume that N is even. The Fast Fourier Transform (FFT) algorithm [37] is used for computing the discrete Fourier transform and enables to speed up the computation if N is a composite number. If N is a prime number, the number of *elementary operations* (a multiplication of two complex numbers followed by addition of two complex numbers) needed for the computation of the discrete Fourier transform is N^2. If $N = N_1 N_2, N_1, N_2 \in \mathbb{N}$, the number of elementary operation can be reduced to $N(N_1 + N_2)$. If $N = N_1 N_2 \dots N_k, N_1, N_2, \dots, N_k \in N$, the number of operations needed is $N(N_1 + N_2 + \cdots + N_k)$. In particular, if $N = 2^k$, it requires $2kN = 2N \log_2 N$ elementary operations [3]. Therefore, we will use N's which are divisible by a higher power of 2, at least 16, better $256, 512$ and are not divisible by a high prime number. Hence it makes no sense to use odd N and in the following text N assume that N is even.

Furthermore, all spectra are shown in logarithmic brightness scale. In linear scale, they would all have a very bright center and no visible structure around it.

Remark 4.9. Digital gray-scale images are a special case of functions discussed in Chapter 3. The discrete Fourier transform of a digital gray-scale image is a complex function whose real and imaginary parts are generally not integer-valued. Therefore, in computers the calculations are performed in floating point types. When exporting the data for visualization, the spectrum or the phase-correlation function or the cross-correlation function is converted to an integer-valued type, usually word (a 16-bit gray-scale image). This



(a) (b)



(c)

Figure 4.2: Original image f (a), its cross-correlation function $Q_{f,f}$ as a gray-scale image (b) and as a 3D graph (c).

way, we loose information on the absolute numbers, which are however of no use for image registration. This is why the values of $Q_{f,f}$ in Figure 4.2c are in $\{0, 1, \ldots, 65\,535\}$ ($65\,535 = 2^{16} - 1$, the image is a 16-bit image).

Remark 4.10. The cross-correlation function of two real functions is again real (Theorems 2.64 on page 34 and 3.27 on page 47). The phase-correlation function and the semi-phase correlation function of two real functions is real too (Theorems 2.66 on page 2.66 and 3.29 on page 48). This enables to draw the correlation functions as usual gray-scale images.

4.2.2 Registration of shifted images

Theoretical results for shifted images are Theorems 2.67 (page 36) and 3.31 (page 49). They both assume that both functions are identical but shifted with respect to each other by a vector (x_0, y_0),

$$f_2(x, y) = f_1(x - x_0, y - y_0).$$

In the discrete case, the support of both functions is only a smaller subset ("subsquare") of $\{0, 1, \ldots, N-1\}^2, N \in \mathbb{N}$ (see Fig. 3.1, page 43). This is a simplified case. In reality, we are trying to find the shift vector between two images which are both cropped from a "bigger scene", in this section from the same bigger image. Here we assume that the images are shifted by integer shifts only ($x_0, y_0 \in \mathbb{Z}$). The phase correlation works very well then. The phase-correlation function is not a discrete impulse function, but there is still one (very clear and sharp) global maximum at coordinates (x_0, y_0). The problem of finding a mutual shift between images is transformed not to finding the non-zero element, but to finding the global maximum.

However, the discrete Fourier transform works either with periodic functions or makes them periodic [13]. In general case, an image has not the same values on the edges and by periodizing an image, we obtain a function with jumps at the edges of the original image. These jumps are often the most contrasty structures in the function and may lead to incorrect registration. Therefore, it is necessary to remove such edges from the image used for the shift estimation, to smooth them out. This is done by multiplying the image by a suitable function g, a so called window function. Such function must be zero or almost zero at the image edges and one on a large part of the image. Commonly used window functions are the Gaussian window function and the Hanning window function. A brief discussion on other types of window functions can be found in [47] and a very detailed discussion can be found in [31].

Definition 4.11 (Window functions). Let sets

$$A = \langle -a, a\rangle \times \langle -b, b\rangle, \qquad a, b \in \mathbb{R}_0^+,$$

$$B = \left\{(x, y); x^2 + y^2 \leq r^2\right\}, \qquad r \in \mathbb{R}_0^+.$$

Let $\sigma \in \mathbb{R}^+$ be a given number. Let $\rho(X, A)$ be the distance of point $X = (x, y)$ from set A, i.e. [18]

$$\rho(X, A) = \inf\left\{d \in \mathbb{R}, d = \rho(X, Y), Y \in A\right\},$$

where $\rho(X, Y)$ is the Euclidean metric.

(a) Function

$$g_{GR}(x, y) = \mathrm{e}^{-\frac{\rho^2(X,A)}{\sigma^2}}$$

is called the *rectangular Gaussian window function.*

(b) Function

$$g_{GC}(x, y) = \mathrm{e}^{-\frac{\rho^2(X,B)}{\sigma^2}}$$

is called the *circular Gaussian window function.*

(c) Function

$$g_{HR}(x,y) = \begin{cases} \frac{1}{2} + \frac{1}{2}\cos\frac{\pi\rho(X,A)}{\sigma} & \text{if } \rho(X,A) \leq \sigma \\ 0 & \text{if } \rho(X,A) > \sigma \end{cases}$$

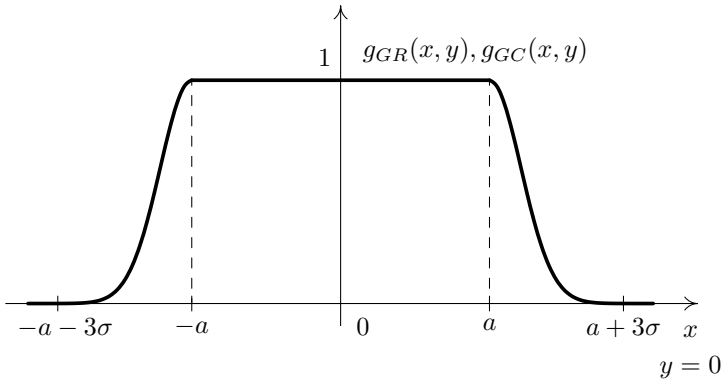
is called the *rectangular Hanning window function.*

(d) Function

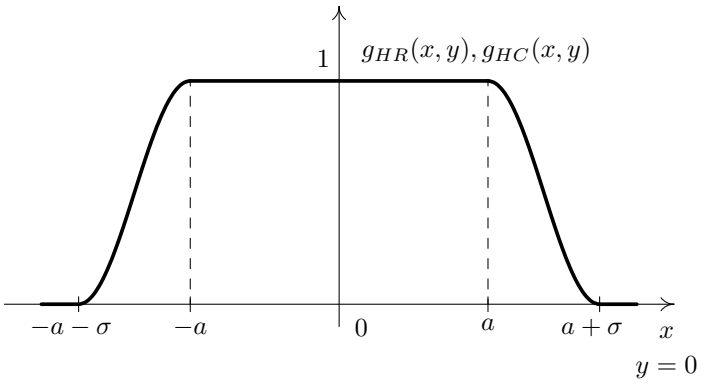
$$g_{HC}(x,y) = \begin{cases} \frac{1}{2} + \frac{1}{2}\cos\frac{\pi\rho(X,B)}{\sigma} & \text{if } \rho(X,B) \leq \sigma \\ 0 & \text{if } \rho(X,B) > \sigma \end{cases}$$

is called the *circular Hanning window function.*

Figure 4.3 shows graphs of cuts of the window functions defined in Definition 4.11. The window functions are defined symmetrical with center in $(0,0)$. When we apply them on images, we must shift them by $\left(\frac{N}{2}, \frac{N}{2}\right)$, i.e. multiply image f by a window function $g\left(x - \frac{N}{2}, y - \frac{N}{2}\right)$.



(a)



(b)

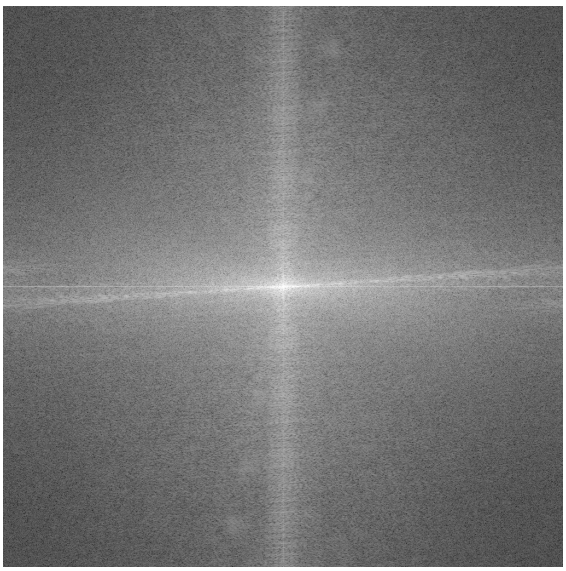
Figure 4.3: Graphs of cut of window functions g_{GR}, g_{GC} (a), g_{HR}, g_{HC} (b) for $y = 0$.

The Gaussian window function never diminishes totally. Therefore, it is necessary to use σ low enough so that the values of g at the image edge are negligible. It is reasonable

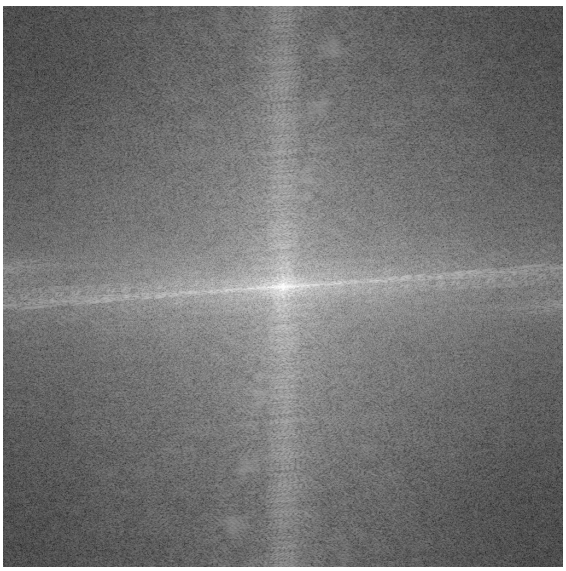
that the distance from the set A or B to the image edge is at least 3σ. For a point (x, y) which is 3σ far from set A, $g_{GR}(x, y) = \mathrm{e}^{-9} \doteq 1.234\,1 \cdot 10^{-4}$, which results in a pixel value of 8 applied on a 16-bit image which is originally white at this pixel. Such value is really negligible compared to other structures in a common image.

The advantage of rectangular window functions is that they keep a bigger part of the image. On the other hand, they still keep some little information on the image edges, as the edges of the unchanged part of the image are parallel with the original image edges. Using circular window functions, we loose a bigger part of the images, but we also loose all information on the original image edges. An advantage of Hanning window functions is that they diminish totally at image edges (for suitable σ). It is not necessary to apply the same window function on both images to be registered. It depends on the distribution of structures in the image. If the window functions are different (i.e. in the best case one is rectangular and the other is circular), it assures that they do not bring about similar structures in the images which might lead to incorrect registration.

The Fourier spectra of most images have a clear cross corresponding to the the jumps at image edges. The purpose of multiplication by a window function is to remove the image edges, which means to remove the cross from the Fourier spectra. Even the rectangular window functions work very well as it is illustrated in Figure 4.4.



(a) Amplitude spectrum of the original image, Fig. 4.2a. Note the slightly inclined horizontal line. It clearly shows that most of the structures in the image are not horizontal, but they are inclined. The boat is not oriented exactly horizontally in the picture.



(b) Amplitude spectrum of the original image multiplied by a rectangular Gaussian window function ($N = 1024$, $a = b = 256$, $\sigma = 85$).

Figure 4.4: Illustration of the image edge removal on the amplitude spectrum.

Even slightly overlapping images can be registered by means of phase correlation. The shift computed by phase-correlation is ambiguous. It is due to the similarity of the Shift Theorem for periodized functions (Thm. 3.16, page 42) and the Shift Theorem 3.17 (page 42). The phase-correlation function does not say if the shift is for instance $\frac{5}{6}N$ or

$-\frac{1}{6}N$. If the global maximum appears at a point close to $(0,0)$, it is natural to assume that the shift is small. If the shift is not small, it is necessary to choose the second possibility. In most cases, it is clearly visible which possibility is the correct one. This makes image registration harder to automatize.

If we register images which are smaller that $N \times N$, we center the images inside the square image. Let us assume that images f_1, f_2 have width m and height n. Let $N \in \mathbb{N}$ be number divisible by a higher power of 2 (at least 16), not divisible by a high prime number, such that $m \leq N, n \leq N$. Then we create image f_{c1}, f_{c2} from images f_1, f_2 by

$$f_{ck}(x,y) = \begin{cases} f_k(x - m_0, y - n_0) & \text{if } m_0 \leq x \leq m_0 + m - 1, n_0 \leq x \leq n_0 + n - 1 \\ 0 & \text{else} \end{cases}$$

for $k = 1, 2$, where $m_0 = \left\lfloor \frac{N-m}{2} \right\rfloor$, $n_0 = \left\lfloor \frac{N-n}{2} \right\rfloor$. Images f_1, f_2 are placed in the center of images f_{c1}, f_{c2} surrounded by a black area. If we center all images into a bigger square one, we can register images of different sizes. Now the edges of the images will be the most contrasty structures in the images and would lead to incorrect registration. Therefore, we must apply suitable window functions on the images to smooth the edges out. It is also possible to apply window functions before centering the images.

4.2.3 Registration of similar images

Generally, similarity is a combination of rotation, scale-change, shift and axial symmetry. It is usually clearly visible whether the transformation between two images involves axial symmetry. If it does, we flip one of the images before proceeding with the registration. Then we can use the method described bellow for estimation of the rotation, scale-change and shift parameters.

Shifts by integer shift vectors can be performed precisely in digital images. Rotations by general angles and scale-changes by general factors must be performed by interpolation of pixel values. Therefore, the Fourier transform of rotated and scaled functions is discussed for functions in $\mathcal{L}(\mathbb{R}^2)$ only (Theorems 2.45 on page 24 and 2.43 on page 23), not in the discrete case.

Let us first consider a digital gray-scale image f_1 and an image f_2, which is a replica of image f_1 rotated by angle $\theta \in \langle 0, 2\pi)$ around the origin, i.e. with some simplification we can write

$$f_2(x,y) = f_1(x\cos\theta - y\sin\theta, x\sin\theta + y\cos\theta).$$

The simplification lies in the fact that the formula treats the functions as if they were $\mathbb{R}^2 \to \mathbb{C}$. Function f_2 must be interpolated and rounded from function f_1 pixel values. If this transformation is performed with digital gray-scale images, a part of image f_1 is cut out in image f_2 and a part is filled with zeroes. In reality, both images are cropped from a "bigger scene", therefore there is no filling with zeroes, pixel values are non-zero even in pixels of f_2 which do not correspond to any pixels in f_1. They are known from the "bigger scene".

Before computing Fourier spectra, we need to apply window functions on both images. If the images are transformed to polar coordinates, rotation is transformed to shift along the angle axis. In most cases, we do not know the point around which the image was rotated, which makes suitable transformation to polar coordinates impossible. According to the Rotation Theorem 2.45, the amplitude spectrum A_2 of image f_2 is rotated by angle

θ with respect to the amplitude spectrum A_1 of image f_1

$$A_2(\xi, \eta) = A_1(\xi \cos\theta - \eta \sin\theta, \xi \sin\theta + \eta \cos\theta).$$

It is always rotated around the origin (from the Similarity Theorem 2.46 on page 25).

Definition 4.12 (Transformation to polar coordinates). Let $f(x, y)$ be a function $\mathbb{R}^2 \to \mathbb{C}$. Function $f(x, y)$ transformed to polar coordinates is function $f^p(\rho, \varphi) : \langle 0, +\infty) \times \langle 0, 2\pi) \to \mathbb{C}$ such that

$$\rho = \sqrt{x^2 + y^2},$$

and φ is the solution to

$$\begin{aligned} x &= \rho \cos\varphi \\ y &= \rho \sin\varphi. \end{aligned}$$

Let us denote

$$A_1^p(\rho, \varphi), A_2^p(\rho, \varphi), \qquad \rho \in \left\langle 0, \frac{N}{2} \right\rangle, \ \varphi \in \langle 0, 2\pi)$$

the amplitude spectra of images f_1, f_2 transformed to polar coordinates. We restrict to $\rho \in \left\langle 0, \frac{N}{2} \right\rangle$, because we assume here that the domain of the image $(0, 1, \ldots, N-1)$ is the same as the domain of the Fourier spectrum and we transform to polar coordinates only the interior of the incircle of the domain. Furthermore the transformed amplitude spectrum is stored in the computer as a digital gray-scale image and it is necessary to decide about the scale of the angle axis. The rotation has changed to shift,

$$\begin{aligned} A_2^p(\rho, \varphi) &= A_1(\rho \cos\varphi \cos\theta - \rho \sin\varphi \sin\theta, \rho \cos\varphi \sin\theta + \rho \sin\varphi \cos\theta) = \\ &= A_1(\rho \cos(\varphi + \theta), \rho \sin(\varphi + \theta)) = A_1^p(\rho, \varphi + \theta), \end{aligned}$$

and we can use phase correlation applied on amplitude spectra A_1^p, A_2^p to determine the shift vector [45]. Yet it is not necessary to use the whole spectra [45]. Corollaries 2.56 (page 31) and 3.20 (page 45) say that if function f is real (as all digital images are)

$$A(\xi, \eta) = A(-\xi, -\eta).$$

Therefore, any two subsequent quadrants of the amplitude spectra contain all information about the whole amplitude spectra. The amplitude spectrum in polar coordinates is a periodic function with period π and we can use the Shift Theorem for Periodized Functions 3.16 (page 42). Applying phase correlation on periodic functions, we do not need to apply any window functions as we do for most digital images. (The amplitude spectra are not periodic in the ρ-axis. Edges in $\rho = 0$ and $\rho = \frac{N}{2}$ may lead to incorrect registration. Since we are not searching for a shift in the ρ-axis, this error is of no interest here.) This means that we apply phase-correlation on $A_1^p(\rho, \varphi), A_2^p(\rho, \varphi), \rho \in \left\langle 0, \frac{N}{2} \right\rangle, \varphi \in \langle 0, \pi)$ and the found shift in angle axis φ represents the rotation angle θ between images f_1 and f_2.

Instead of functions $A_1^p(\rho, \varphi), A_2^p(\rho, \varphi)$, it is better to use functions $\ln(1 + A_1^p(\rho, \varphi))$, $\ln(1 + A_2^p(\rho, \varphi))$. Amplitude spectra A_1, A_2 have the highest values around $(0, 0)$ and higher spatial frequencies have much smaller values of the amplitude spectra. The lowest frequencies are not that useful for phase-correlation. They describe the overall brightness of the image, vignetting, diffuse light in the optical system etc. Applying logarithm on

the amplitude spectra values makes higher spatial frequencies more significant for phase correlation.

The rotation angle computed by phase correlation is ambiguous. The phase correlation does not say if the rotation angle is e.g. $\frac{2}{3}\pi$ or $\frac{5}{3}\pi$ [45]. It is due to the π-periodicity of amplitude spectra A_1^p, A_2^p. It is usually reasonable to await small rotation angles and if they are not correct, to choose the second possibility. This makes registration of rotated images harder to automatize. [45] suggest a method to automatize it. First, f_2 is rotated by the computed angle θ, then by $\pi + \theta$. In both cases, the shift vector (x_0, y_0) between images f_1 and f_2 rotated by $-\theta$ is computed by means of phase correlation. The higher peak in the phase-correlation function corresponds with the correct rotation.

Remark 4.13. De Castro and Morandi in [16] described a method for phase-correlation based registration of translated and rotated images. They introduced a function

$$G(\xi, \eta, \vartheta) = \frac{F_2(\xi, \eta)}{F_1(\xi \cos\vartheta - \eta \sin\vartheta, \xi \sin\vartheta + \eta \cos\vartheta)},$$

where $\vartheta \in \langle 0, 2\pi)$ is a variable. Obviously, if $\vartheta = \theta$,

$$G(\xi, \eta, \theta) = \mathrm{e}^{-\mathrm{i}(\xi x_0 + \eta y_0)}$$

and the inverse Fourier transform of $G(\xi, \eta, \theta)$ is the Dirac distribution shifted by $(-x_0, -y_0)$. The "closer" $\mathcal{F}^{-1}\{G(\xi, \eta, \vartheta)\}$ is to the Dirac distribution (or the discrete impulse function in the practical computation) the better estimation of θ is ϑ. We must evaluate $\mathcal{F}^{-1}\{G(\xi, \eta, \vartheta)\}$ for a number of values of ϑ and then iterate in the interval which gives the best estimations. Each evaluation requires one computation of the inverse Fourier transform, whereas the method based on amplitude spectra which was discussed above in this section requires four Fourier transforms and one inverse Fourier transform in total. This makes the method of De Castro and Morandi significantly slower than the method based on amplitude spectra.

Let us now consider image f_2 not only rotated by angle θ, but also scaled by factor $\alpha \in \mathbb{R}^+$ with respect to image f_1, i.e. with simplification we can write

$$f_2(x, y) = f_1(\alpha x \cos\theta - \alpha y \sin\theta, \alpha x \sin\theta + \alpha y \cos\theta).$$

Definition 4.14 (Transformation to logarithmic-polar coordinates). Let $f(x, y)$ be a function $\mathbb{R}^2 \to \mathbb{C}$. Function $f(x, y)$ transformed to logarithmic-polar coordinates is function $f^{lp}(\rho, \varphi) : \mathbb{R} \times \langle 0, 2\pi) \to \mathbb{C}$ such that for

$$\mathrm{e}^{\rho} = \sqrt{x^2 + y^2},$$

and φ is the solution to

$$\begin{aligned} x &= \mathrm{e}^{\rho} \cos\varphi \\ y &= \mathrm{e}^{\rho} \sin\varphi. \end{aligned}$$

Before computing Fourier spectra, we apply window functions on both images. According to the Similarity Theorem 2.46 (page 25), the amplitude spectrum of image f_2 is

(up to a multiplicative constant) rotated by angle θ and scaled by factor $\frac{1}{\alpha}$ with respect to amplitude spectrum A_1,

$$A_2(\xi,\eta) = \frac{1}{\alpha^2} A_1\left(\frac{\xi}{\alpha}\cos\theta - \frac{\eta}{\alpha}\sin\theta, \frac{\xi}{\alpha}\sin\theta + \frac{\eta}{\alpha}\cos\theta\right).$$

Let us denote

$$A_1^{pl}(\rho,\varphi), A_2^{pl}(\rho,\varphi), \qquad \rho \in \left(0, \ln\frac{N}{2}\right\rangle, \varphi \in \langle 0,\pi)$$

amplitude spectra A_1, A_2 transformed to logarithmic-polar coordinates. The lower bound for ρ was set to 0, which leaves out the value at the origin. This is the lowest spatial frequency, whose value only says the average image brightness. This number has no sense for image registration. We are taking $\varphi \in \langle 0,\pi)$ only as the amplitude spectra are π-periodic (due to Corollaries 2.57 on page 31 and 3.21 on page 45). Changing the linear scale on the ρ axis to logarithmic undersamples the functions, therefore it is necessary to decide well about the scale of the ρ axis.

After the amplitude spectra were transformed to logarithmic-polar coordinates, the rescale changed to shift on the ρ axis,

$$\begin{aligned} A_2^{lp}(\rho,\varphi) &= \frac{1}{\alpha^2} A_1\left(\frac{\mathrm{e}^\rho}{\alpha}\cos\varphi\cos\theta - \frac{\mathrm{e}^\rho}{\alpha}\sin\varphi\sin\theta, \frac{\mathrm{e}^\rho}{\alpha}\cos\varphi\sin\theta + \frac{\mathrm{e}^\rho}{\alpha}\sin\varphi\cos\theta\right) = \\ &= A_1\left(\mathrm{e}^{\rho-\ln\alpha}\cos(\varphi+\theta), \mathrm{e}^{\rho-\ln\alpha}\sin(\varphi+\theta)\right) = A_1^{lp}(\rho-\ln\alpha, \varphi+\theta), \end{aligned}$$

and we can apply phase correlation to the logarithmic-polar amplitude spectra A_1^{lp}, A_2^{lp} [45], [1]. The found shift vector represents the logarithm of the scale factor, $\ln\alpha$ on the ρ-axis and the rotation angle θ in the φ-axis. Similarly to registration of rotated images, it is better to use functions $\ln(1 + A_1^{lp}(\rho,\varphi)), \ln(1 + A_2^{lp}(\rho,\varphi))$ instead of A_1^{lp}, A_2^{lp} for the phase correlation. Furthermore, it is necessary to apply a rectangular window function on the amplitude spectra, since they are not periodic in the ρ-axis. Constants a and b of the window function should be chosen so that only edges with $\rho = 0$ and $\rho = \ln\frac{N}{2}$ are removed. The amplitude spectra are π-periodic, therefore there is not need to remove the edges with $\varphi = 0$ and $\varphi = \pi$. Influences of image rotation and scale-change on amplitude spectra and amplitude spectra in logarithmic-polar coordinates are illustrated in Figure 4.5

If image f_2 is not only a rotated and scaled, but also a shifted replica of image f_1 (shift and rotation is equivalent with rotation around an arbitrary point),

$$f_2(x,y) = f_1(\alpha x\cos\theta - \alpha y\sin\theta - x_0, \alpha x\sin\theta + \alpha y\cos\theta - y_0),$$

where $x_0, y_0 \in \mathbb{Z}, |x_0|, |y_0| < N$, the Similarity Transformation Theorem 2.46 (page 25) implies that the shift has no influence on the amplitude spectra of image f_2,

$$A_2(\xi,\eta) = \frac{1}{\alpha^2} A_1\left(\frac{\xi}{\alpha}\cos\theta - \frac{\eta}{\alpha}\sin\theta, \frac{\xi}{\alpha}\sin\theta + \frac{\eta}{\alpha}\cos\theta\right).$$

Therefore, we can find the angle θ and scale factor α by means of the method described above. After image f_2 is rotated by angle $-\theta$ and scaled by factor $\frac{1}{\alpha}$ to compensate the rotation and scale-change found by the phase correlation, creating image f_3, only shift remains between image f_1 and image f_3. Now we can apply phase correlation to find the shift [45].

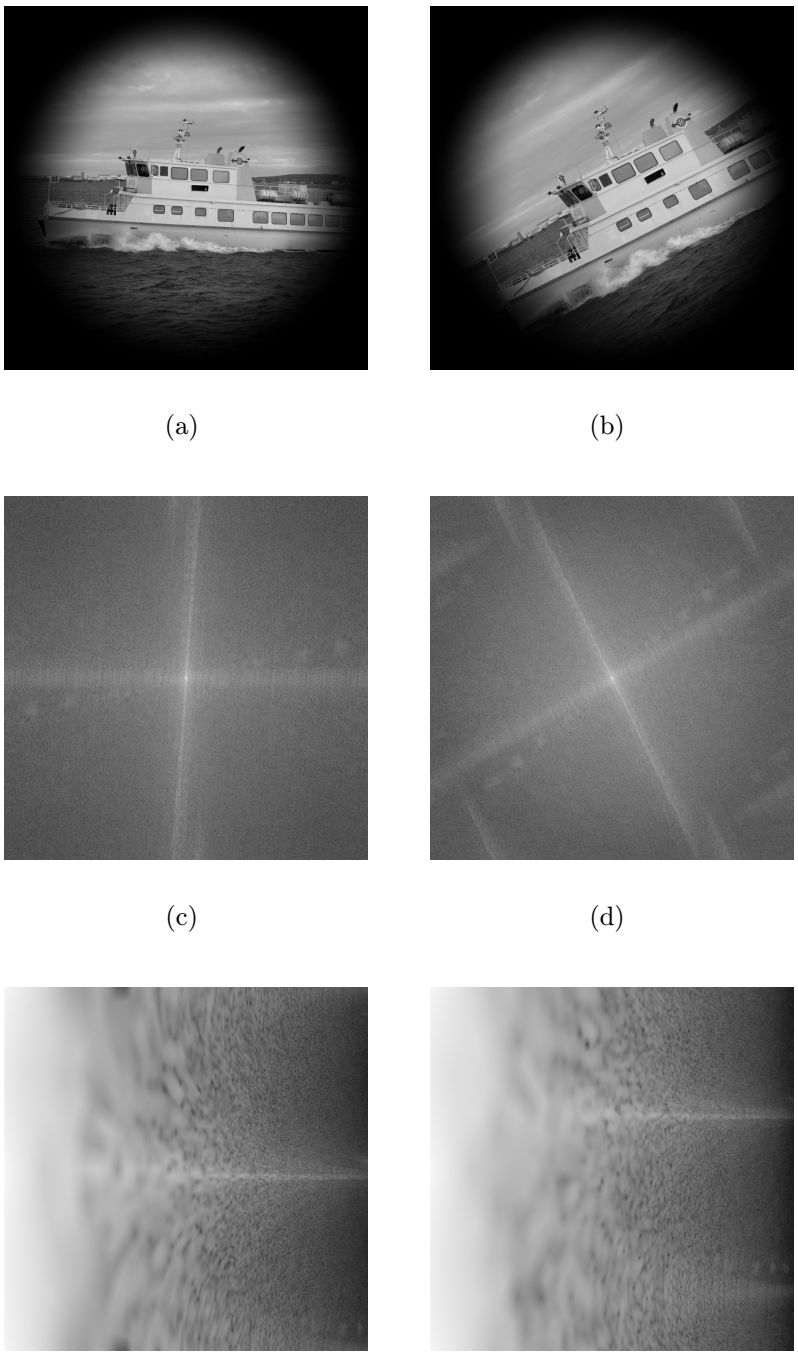


Figure 4.5: Illustration of influences of image rotation and scale on amplitude spectra and amplitude spectra in logarithmic-polar coordinates. Images (a) and (b) are both made from the same original image, image (b) is cropped from an image rotated by 30 degrees and scaled by $\alpha = 0.8$ with respect to the original one. A circular Hanning window function was applied on both images. Images (c) and (d) show amplitude spectra of images (a) and (b). Images (e) and (f) show the amplitude spectra transformed to logarithmic-polar coordinates.

Influence of matrix size N on registration precision

When we considered shifted images with integer shifts, it was enough that the size of the matrices representing the Fourier spectra was the same as the size of a square that the digital images fitted in. The shift was found precisely by phase correlation. If we consider rotated and scaled images, the situation is totally different. Apart from rotation by $k\frac{\pi}{2}, k \in \mathbb{Z}$, there is nothing like "integer rotation". The size of the amplitude-spectrum matrix in logarithmic-polar coordinates says how many steps the angle of π is divided to. We can find the rotation between two images up to this precision only (if we do not consider the subpixel extension of the phase-correlation method discussed in Section 4.4 on page 74). Similarly, the precision of the registration of scaled images depends on the resolution of the ρ-axis.

4.3 Registration of real images

In this section, we assume that the images to be registered are not identical up to a similarity. They may be taken with different exposure settings, there may be dust particles on the chip. The images may be taken by different cameras causing different properties of additive and impulse noise and through different optical systems with different modulation transfer functions.

The *modulation transfer function* [21] of an image is a function $\kappa = k(\tau)$, where τ is spatial frequency and k the normalized contrast. A testing image is a pattern composed of black and white stripes with sine brightness course. The spatial frequency is the number of stripe couples in a unit length. Thus the spatial frequency is measured in m^{-1}, usually in mm^{-1} (often line pairs per millimeter (lp/mm) are used). The normalized contrast is the difference between the highest and the lowest pixel values in the image of the stripes divided by the highest and the lowest possible pixel values, i.e.

$$k = \frac{S - s}{T - t},$$

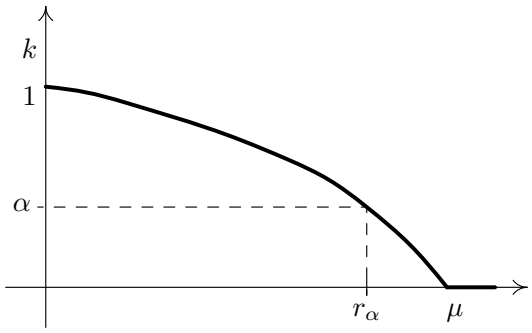
where S and s are the maximal and minimal pixel values in an image of the stripes with spatial frequency τ, T and t are the theoretical maximal and minimal pixel values. The lowest value of $\tau > 0$ for which $k(\tau) = 0$ is denoted by μ and called the *limiting frequency*. For a given normalized contrast α, the frequency $r_\alpha = \kappa^{-1}(\alpha)$ is called the *resolution at contrast* α.

Examples of the modulation transfer function are in Figure 4.6.

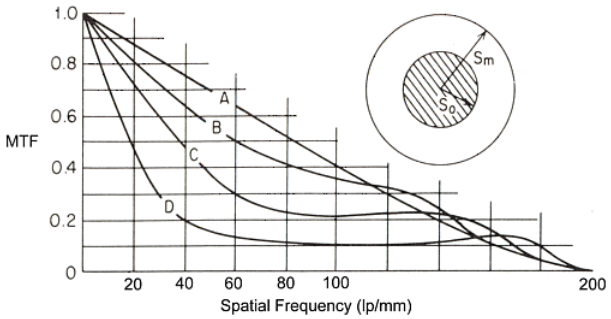
Definition 4.15 (Image defects, impulse noise). Let f be a digital gray-scale image representing an ideal image. Let h be a digital gray-scale image with pixel values

$$h(x,y) = \begin{cases} f(x,y) & \text{if } Y(x,y) = 0, \\ m(x,y) & \text{if } Y(x,y) = 1, \end{cases}$$

where $Y(x,y)$ are realizations of a random variable X with Bernoulli distribution and $m(x,y)$ are independent realization of a random variable Z (generally, we do not know its distribution). If $Y(x,y)$ are not independent realizations of X, we say that image h contains *defects*. If $Y(x,y)$ are independent realizations of X, we say that image h contains *impulse noise* [22]. The impulse noise is called *constant* if for the same conditions (images



(a) An model example of a graph of a modulation transfer function.



(b) The modulation transfer function of the Canon RF500/4.5L lens at f8 (curve A). Curves B, C and D apply for mirror lenses with central obstruction (sketched above the graph) of the same parameters as the Canon lens. The ratio $S_0 : S_m$ is 0.25 (curve B), 0.5 (curve C), and 0.75 (curve D). Graph taken from [5].

Figure 4.6: Examples of the modulation transfer functions.

taken with the same camera and same exposure setting shortly one after another), we obtain the same realization of X. The impulse noise is called *variable* if the realizations of X are different for every image.

Remark 4.16. Similarly to additive noise, we need to treat the case when $f+m > w-1$ or $f+m < 0$ ($w-1$ is the maximal possible pixel value, see Def. 4.1). On the contrary to the additive noise, the exposure setting is not adjusted to impulse noise. All values higher than $w-1$ are mapped to $w-1$ and values lower than 0 are mapped to 0 like we did for additive noise in Remark 4.5 on page 51.

The cause of image defects and constant impulse noise in digital images are mostly faulty sensors. In scanned classical images, image defects are for example scratches on the scanned film. The cause of variable impulse noise in digital images are mostly collisions of the chip with high-energy particles (cosmical radiation). They result in saturated ($h(x,y) = w-1$) pixels called the *hot pixels*. The longer exposure time we use, the more hot pixels the image contains.

Definition 4.17 (Dust). Let f be a digital gray-scale image representing an ideal image containing no dust. Let s be a function $\mathbb{R}^2 \to \langle 0,1\rangle$ which is equal to 1 on $\mathbb{R}^2$ apart from a finite number of bounded connected sets $A_1, \dots A_n, n \in \mathbb{N}$. Let image

$$h(x,y) = \mathrm{Round}(f(x,y)s(x,y)).$$

Then we say that image h contains *dust*.

Remark 4.18. In digital cameras, there may be dust particles directly on the chip or more often on a filter before it. If the dust particles are directly on the chip, it causes almost black spots in the image which are almost independent on the used optical system. If the dust particles are on a filter, they cause darker spots in the image whose size and dimness depend on the optical system and the aperture setting (if the system has an

aperture). Therefore, the appearance of dust particles may vary from image to image. Furthermore, if we shake the camera or the camera shakes the chip (some newer camera models shake the chip to remove dust particles), the particles may move. If we do not change the settings of the optical system and do not move the camera much, we can take so called *flat-field* images, out-of-focus images of an evenly lit ground (e.g. the sky through a flimsy paper). They enable us to estimate the function s and if we multiply function h by $\frac{1}{s}$, we can somehow restore image f provided s is not 0 or close to zero at some pixels. However, further image processing may reveal the places with corrected darker spots again.

4.3.1 Registration of images with respect to their modulation transfer function

The modulation transfer function is mainly influenced by the properties of the optical system and the photo-sensitive elements of the chip. If the optical system is out of focus, the limiting frequency (from the definition of the modulation transfer function on page 63) is lower than if the optical system is in focus. If we have a very good lens and a chip with large sensors, the limiting frequency is lower than if we had smaller sensors. Taking the size of sensors into consideration, the limiting frequency can be used for computing the limiting frequency in the Fourier spectrum of the image taken with this system. If we have focused optics of high quality enough for the size of the sensors, the limiting frequency in the Fourier spectrum is larger than $\frac{\sqrt{2}}{2}N$ (the highest frequency of the spectrum). If the optical system is out of focus, some higher frequencies in the Fourier spectrum may contain no information, be equal to zero (or almost zero if we take noise into consideration). Such case can be treated by using the semi-phase correlation function (Definition 3.25 on page 46)

$$P^{p,q}_{f_1,f_2}(x,y) = \mathcal{D}^{-1}\left\{\frac{F_1(\xi,\eta)\cdot F_2^*(\xi,\eta)}{(|F_1(\xi,\eta)|+p)\cdot(|F_2(\xi,\eta)|+q)}\right\}$$

for $p,q \in \mathbb{R}^+$. Letting $(p,q) \to (0,0)$, we obtain the phase-correlation function. If we take high values of p,q, we are dividing by almost a constant, resulting in nearly the cross-correlation function (up to a multiplicative constant). If we use pure phase-correlation, the values of the cross-power spectrum are normalized by its amplitude, which means that information on all spatial frequencies are taken with the same weight. If there is almost no information on a given frequency, it is divided by its very small amplitude. If there is much information on a given spatial frequency, the value of the cross-power spectrum is large causing division by a large amplitude. On the contrary, a pure cross-correlation does not perform any normalization, the cross-power spectrum is taken as it is. This may seem to be better in case there are frequencies with no information, like in an image out of focus. However, as wee have seen in Figure 4.2 on page 54, the cross-correlation function does not necessarily have the global maximum as a very sharp and high peak. It may be instable for registration of images with different properties. Therefore, it is better to use a semi-phase correlation. For sharp images, we may use very small values of p,q only to make sure that no division by zero may occur. For images which are out-of-focus, taken with poor optics, or at bad conditions (seeing, haze), we should increase p and q to decrease the influence of frequencies with no information. Suitable values of p,q are $0.01\max_{\xi,\eta\in\{0,1,N-1\}} A_f(\xi,\eta)$ to $0.1\max_{\xi,\eta\in\{0,1,N-1\}} A_f(\xi,\eta)$ (where A_f is the amplitude

spectrum of function f) [23]. If the two images to be registered were taken with the same optical system, both constants should be the same. However, to register images taken with optical system of very different qualities, we can use different values of p, q.

[6] uses the semi-phase correlation function also for registration of images deteriorated by Gaussian additive noise.

4.3.2 Registration of images containing additive noise and variable impulse noise, with vignetting and diffuse light

Both additive noise and variable impulse noise are different in every image. Therefore, there are no corresponding structures in the noise between two different images to be registered. Both additive noise and variable impulse noise represent information on all spatial frequencies, mostly on the highest ones. Usually, they are dominant on these frequencies, which makes the highest frequencies less reliable for image registration. Furthermore, the lowest spatial frequencies are useless too. $F(0,0)$ represents the sum of values of all pixels (Theorem 3.13) (in a gray-scale image the average pixel value) and other low frequencies contain information about the largest-scale structures in the images such as optics vignetting and diffuse light in the optical system. They are sensitive to different contrast and brightness settings of the images. Removing the lowest spatial frequencies is also a way of enhancing the image structure and is used for visualizing images with high dynamic range, which makes it suitable for registration of such images, especially with structures which are hardly visible in the image. A modification of this method is also used for images of the solar corona discussed in Section 4.6 (starting from page 75).

Due to the facts mentioned above, it is often necessary to multiply the Fourier spectra of the images by a suitable weight function to reduce information on the highest and the lowest spatial frequencies. Multiplying the Fourier spectra by a low-pass high-pass weight function, the semi-phase correlation function (for any $p, q \geq 0$) with this modification remains real, which is ensured by Corollary 3.21 on page 45.

Definition 4.19 (Low-pass high-pass weight function). Let $r_1, r_2, \sigma_1, \sigma_2 \in \mathbb{R}^+$ such that $r_1 < r_2$. Function $H_{r_1,\sigma_1}(\xi,\eta) : \mathbb{R}^2 \to \langle 0,1\rangle$ defined as

$$H_{r_1,\sigma_1}(\xi,\eta) = \begin{cases} 0 & \text{if } \frac{4}{N^2}(\xi^2+\eta^2) < (r_1-\sigma_1)^2 \\ \frac{1}{2} + \frac{1}{2}\cos\frac{\pi\left(r_1 - \frac{2}{N}\sqrt{\xi^2+\eta^2}\right)}{\sigma_1} & \text{if } (r_1-\sigma_1)^2 \leq \frac{4}{N^2}(\xi^2+\eta^2) < r_1^2 \\ 1 & \text{else} \end{cases}$$

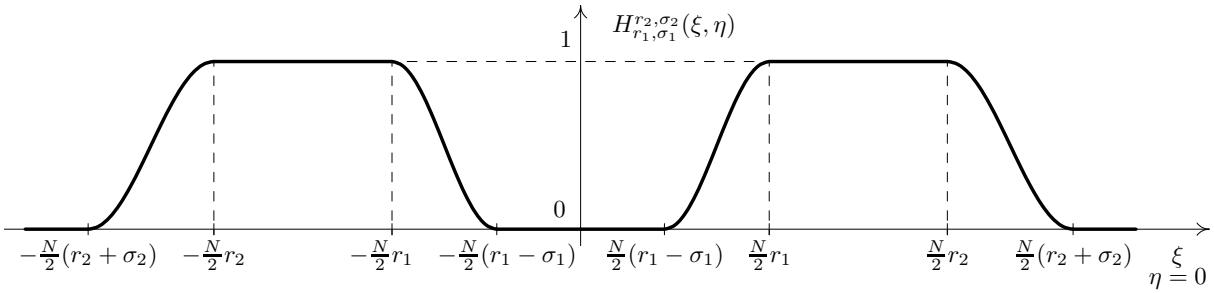
is called the *high-pass weight function*. Function $H^{r_2,\sigma_2}(\xi,\eta) : \mathbb{R}^2 \to \langle 0,1\rangle$ defined as

$$H^{r_2,\sigma_2}(\xi,\eta) = \begin{cases} 1 & \text{if } \frac{4}{N^2}(\xi^2+\eta^2) < r_2^2 \\ \frac{1}{2} + \frac{1}{2}\cos\frac{\pi\left(r_2 - \frac{2}{N}\sqrt{\xi^2+\eta^2}\right)}{\sigma_2} & \text{if } r_2^2 \leq \frac{4}{N^2}(\xi^2+\eta^2) < (r_2+\sigma_2)^2 \\ 0 & \text{else} \end{cases}$$

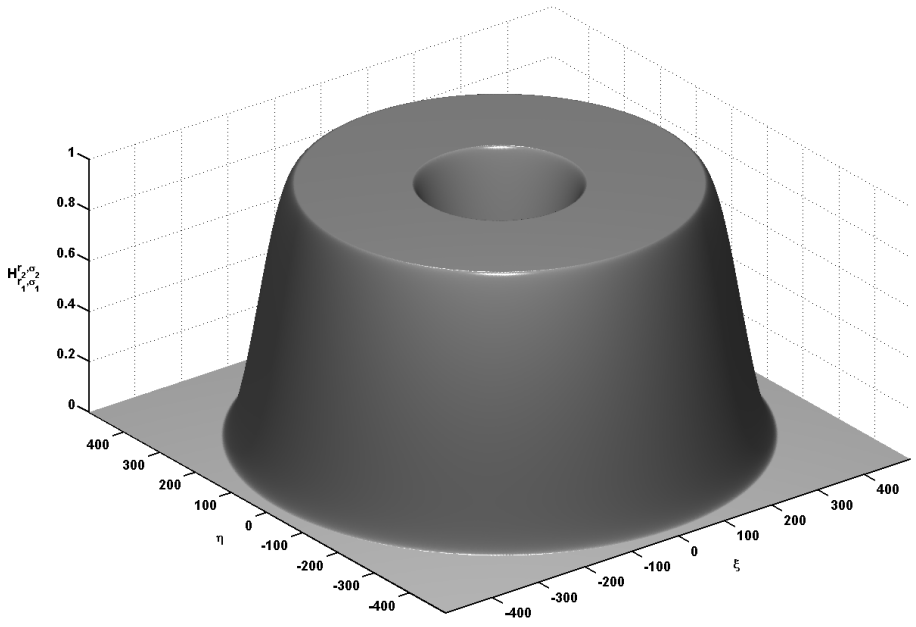
is called the *low-pass weight function*. Function $H^{r_2,\sigma_2}_{r_1,\sigma_1}(\xi,\eta) : \mathbb{R}^2 \to \langle 0,1\rangle$ defined as

$$H^{r_2,\sigma_2}_{r_1,\sigma_1}(\xi,\eta) = H_{r_1,\sigma_1}(\xi,\eta) \cdot H^{r_2,\sigma_2}(\xi,\eta)$$

is called the *low-pass high-pass weight function*. Graphs of a low-pass high-pass weight function are in Figure 4.7.



(a) Graph of cut of function $H_{r_1,\sigma_1}^{r_2,\sigma_2}(\xi,\eta)$.



(b) Graph of function $H_{r_1,\sigma_1}^{r_2,\sigma_2}(\xi,\eta)$ with $N = 1\,000$, $r_1 = 0.3$, $r_2 = 0.7$, $\sigma_1 = 0.2$, $\sigma_2 = 0.25$.

Figure 4.7: Graphs of a low-pass high-pass weight function.

A similar weight function is used in [45]. Like window functions, the weight functions from Definition 4.19 are defined symmetrical with center in $(0, 0)$. When we apply them on images, we must shift them by $\left(\frac{N}{2}, \frac{N}{2}\right)$, i.e. multiply the Fourier spectra of the images by function $H_{r_1,\sigma_1}^{r_2,\sigma_2}\left(\xi - \frac{N}{2}, \eta - \frac{N}{2}\right)$.

Multiplying the spectra by a low-pass high-pass weight function causes that the product is equal to zero for some values of ξ and η. Therefore, it is necessary to use the semi-phase correlation with $p, q > 0$ to avoid division by zero. In a general case, it is possible to apply different low-pass high-pass weight functions on each image to be registered. For instance, additive noise in an image scanned from a classical film has different properties than the noise in an image taken with a digital camera. However, choosing appropriate values of $r_1, r_2, \sigma_1, \sigma_2$ is complicated and there is no precise rule for that. It is usually necessary to set the parameters manually, study the modified phase-correlation function and if the shift vector is not estimated correctly, change the them and repeat the

computation.

In most cases, we do not make the image registration less precise if we apply one low-pass high-pass weight function on the normalized cross-power spectrum of the images. It enables to recompute the modified phase-correlation function faster after having changed the parameters $r_1, r_2, \sigma_1, \sigma_2$ and it involves less parameters. It also does not bring about division by zero, because we multiply by the weight function after having divided. In the further text, we will stick to this simpler case. Then the formula for modified phase correlation for images with additive noise and variable impulse noise is

$$\mathcal{D}^{-1}\left\{H_{r_1,\sigma_1}^{r_2,\sigma_2}\left(\frac{F_1(\xi,\eta)\cdot F_2^*(\xi,\eta)}{(|F_1(\xi,\eta)|+p)\cdot(|F_2(\xi,\eta)|+q)}\right)\right\}$$

with $p, q \geq 0$ and the result of the inverse discrete Fourier transform remains real due to Corollary 3.21 (page 45).

If there is no need for the high-pass weight function due to negligible amount of additive noise and variable impulse noise or the application of a high-pass weight function is performed in another way (for example by a filter applied directly on images f_1, f_2 as it is described in Section 4.6 starting from page 75), the low-pass weight function from Definition 4.19 can be replaced by the Gaussian low-pass weight function, which treats the normalized cross-power spectrum in a different way.

Definition 4.20 (Gaussian low-pass weight function). Let $\lambda \in \mathbb{R}_0^+$. Function $H_\lambda(\xi,\eta) : \mathbb{R}^2 \to \langle 0,1\rangle$ defined as

$$H_\lambda(\xi,\eta) = \mathrm{e}^{-\lambda\frac{\xi^2+\eta^2}{N^2}}$$

is called the *Gaussian low-pass weight function* with parameter λ.

In fact, the Gaussian low-pass weight function is identical with the circular Gaussian window function from Definition 4.11 (page 55) for $r = 0$ and $\lambda = \frac{N^2}{\sigma^2}$. Influence of the λ parameter of the Gaussian low-pass weight function is illustrated in Figures 4.8 and 4.9.

4.3.3 Registration of images containing image defects, constant impulse noise and dust

In this section, we will assume that the images to be registered contain not only additive noise and variable impulse noise, but also image defects, constant impulse noise and dust and both of them were taken with the same camera. Images taken by different cameras have faulty pixels and dust at different places. Then these structures in one image do not correspond to any structures in the second image and they do not bring about a misleading maximum of the modified phase-correlation function. The difference between the noise discussed in the previous chapter and the defects discussed here is that image defects, constant impulse noise and dust represent structures which are identical or very similar in both images. Therefore, they often lead to incorrect image registration. The shift computed by phase-correlation may be $(0,0)$, because the faulty pixels and dust do not move between taking images. This causes the worst problems if the shift between the images is very small and the correct maximum is then close to $(0,0)$. If the norm of the shift vector is higher (the maxima corresponding to the correct shift and the zero shift are far enough so that their peaks do not overlap), the maxima are clearly distinguishable and

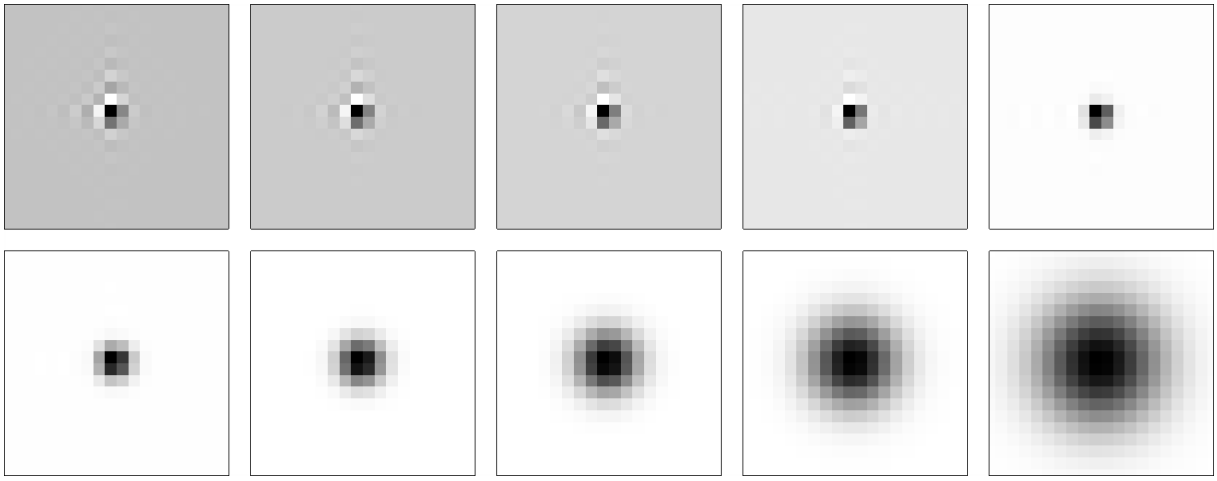


Figure 4.8: Details of peaks of the phase-correlation function with Gaussian low-pass weight function applied on the normalized cross-power spectrum. Parameter λ is set to $0, 1, 2, 4, 8, \ldots, 256$ (from top left to bottom right). The images are shown in negative, i.e. the darkest pixels are global maxima. Image f_1 is a high quality image of the solar corona taken by Constantinos Emmanoulidis in Akademgorodok, Novosibirsk, Russia on August 1, 2008 with digital camera Canon EOS 350D and Skywatcher ED80 lens with f6.3 focal reducer 395 mm. Image f_2 is identical with image f_1 up to a non-integer shift by $(5.352, -3.598)$. A circular Gaussian window function and the T_8 filter were applied on both images (see Section 4.6.1 for details on the filter). The details show 20×20 pixels. $\lambda = 32$ is probably the best for estimation of the shift vector.

by choosing suitable values of the parameters of the low-pass high-pass weight function we can usually assure that global maximum is the correct one. This may sometimes deteriorate the registration precision.

Even if the images contain image defects, constant impulse noise and dust, we may try to register them without treating the defects. Treating them takes more computing time and in case of dust, it cannot be done automatically. It often helps to adjust the value of r_1 or λ of the low-pass weight function applied on the normalized cross-power spectrum. However, if the peak of the modified phase-correlation function becomes too flat and the registration imprecise, it is necessary to treat the image defects, constant impulse noise or dust. If we treat all of them, it is necessary to treat image defects and impulse noise before treating dust, so that we do not use faulty pixels for interpolation of spots caused by dust particles. Dust causes bigger problems than constant impulse noise. Impulse noise affects single pixels, having impact especially high spatial frequencies, which are then filtered out by the low-pass weight function. The spots from dust particles are bigger, affecting more frequencies.

Treating image defects and impulse noise

Treating impulse noise was described in [22]. It always consists of two steps — detection and correction.

To detect impulse noise and image defects, we test a statistical hypothesis that a given

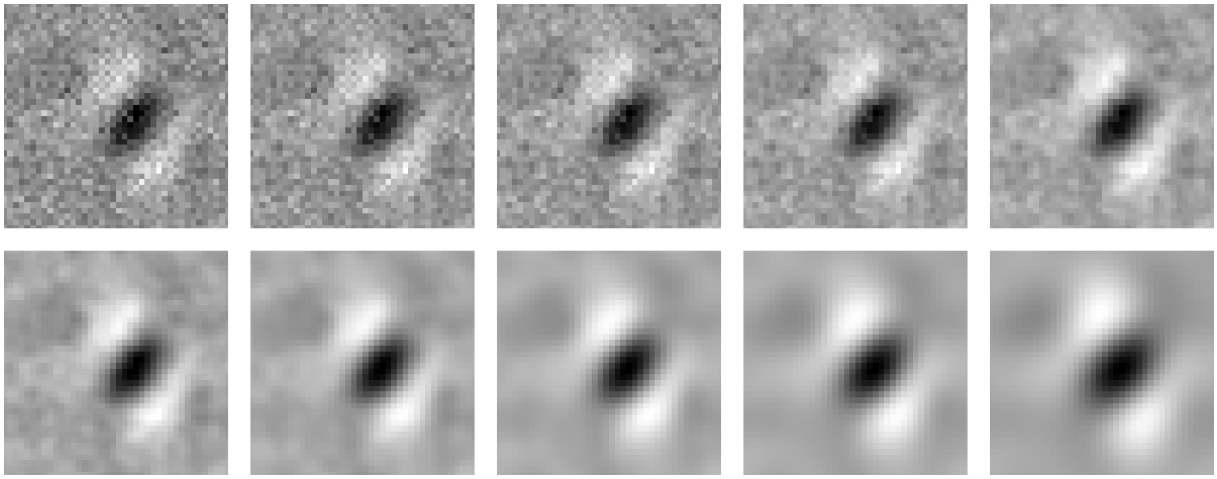


Figure 4.9: Details of peaks of the phase-correlation function with Gaussian low-pass weight function applied on the normalized cross-power spectrum. Parameter λ is set to $0, 1, 2, 4, 8, \ldots, 256$ (from top left to bottom right). The images are shown in negative, i.e. the darkest pixels are global maxima. Image f_2 is a slightly blurred image (the lens was out of focus, there was probably motion-blur too) of the solar corona taken by Lapo Casetti at the Bodonch Gol in Mongolia on August 1, 2008 with digital camera Canon EOS 350D and Borg 77ED refractor, 6.5/500 mm, with exposure time 2 s. Image f_1 is a composition of ten images with exposure times 1/125 s to 1/4 s. A circular Gaussian window function and the T_{10} filter were applied on both images. The details show 40×40 pixels. $\lambda = 64$ is probably the best for estimation of the shift vector.

pixel (x, y) is faulty. The test criteria is

$$|f(x, y) - \bar{f}(x, y)| < \epsilon,$$

where $f(x, y)$ is the value of the given pixel in the tested image and $\bar{f}(x, y)$ is a location characteristic of the statistical set consisting of pixel values of the pixels in the neighborhood of pixel (x, y). A very good characteristic to be used is median though it is slower to be computed. It is faster to compute the mean value. The mean value gives worse results especially in case there are more faulty pixels in the neighborhood of the tested pixel, e.g. a row of faulty pixels. Probably the best characteristic is the trimmed mean value, i.e. a mean value computed from only those pixel values $f(s, t)$ of the neighborhood which have

$$|f(s, t) - f_{0.5}(x, y)| < \delta,$$

where $\delta > \epsilon$ and $f_{0.5}(x, y)$ is the median of pixel values in the neighborhood of pixel (x, y). It is convenient for computations to use square neighborhoods, e.g. 5×5 pixels. There is a mathematical method for choosing the appropriate values of ϵ described in [22], but it is usually enough to check visually if the faulty pixels are removed and the image is not deteriorated much.

A fast single-path method of faulty-pixel correction replaces the pixel value $f(x, y)$ by the value of the location characteristic used for detection, $\bar{f}(x, y)$. In this case, the value $\bar{f}(x, y)$ is computed from all pixels, even from the faulty ones. This gives bad results especially if we use the mean value as the location characteristic. A slower, double-path method replaces $f(x, y)$ by $\bar{f}(x, y)$ computed only from pixels which were not detected as

faulty. The best and the slowest method replaces $f(x, y)$ by a value interpolated from the neighboring pixels which were not detected as faulty.

If there are bigger image defects, e.g. a block of faulty pixels or multiple faulty rows, they need to be detected manually and then interpolated from the neighboring pixels.

Treating dust

Darker spots in the images caused by dust particles on the chip or on the filters before it are several pixels to tens of pixels large. They are too large to be detected by a method used for detecting impulse noise and may look similar to correct structures in the image. Therefore, they must be detected manually. After selecting which pixels belong to the spots, the pixels should be interpolated from the surrounding pixels. The interpolation should be done in a way that does not bring about any new structures in the affected pixels. A suitable method is incomplete convolution described in [25].

4.3.4 Registration of shifted images step by step

In this section, we will sum up the modified phase correlation for shifted images. Let us denote f_1, f_2 the images to be registered. For computing the Fourier transform we will use a square $N \times N$ (N is an even number).

1. Multiplication of images f_1, f_2 by window functions, obtaining images f_{w1}, f_{w2}.
2. Centering images f_{w1}, f_{w2} in the square $N \times N$ pixels, obtaining images f_{c1}, f_{c2}.
3. Computing the semi-normalized cross-power spectrum $Z = Z^{p,q}_{f_{c1},f_{c2}}$ of images f_{c1}, f_{c2}. Constants p, q chosen close to zero.
4. Multiplication of Z by a low-pass high-pass weight function, obtaining function Z_w.
5. Computing the inverse discrete Fourier transform P of Z_w.
6. Finding the coordinates of the global maximum of function P, i.e. finding the shift vector $(-x_0, -y_0)$.

The algorithm involves two discrete Fourier transforms and one inverse discrete Fourier transform computed by means of the Fast Fourier Transform algorithm.

Applying the computed shift on the original images and checking their coincidence, we may find out that the computed shift is incorrect. From the properties of images f_1, f_2 and the shape of the peak in P, we can suggest changes to the algorithm to compute a better shift vector. The fastest option is to change the properties of the low-pass high-pass weight function. Enhancing parameter r_1 or λ usually helps in cases the global maximum is not the centroid of the peak and also in cases when the global maximum is in $(0, 0)$ (corresponding to the dust and constant impulse noise). In the latter case, it would be better to remove the dust or constant impulse noise, but it is more time-consuming (dust must be removed manually and if we remove constant impulse noise, we must start the algorithm from scratch). If we do not need to enhance r_1 or λ too much ($r_1 = 0.5$ or $\lambda = 512$ is about the maximal usable value), it is possible to overcome the problem of dust and constant impulse noise by the low-pass weight function. If there is too much dust or constant impulse noise, it is necessary to treat it. Constant impulse noise must be treated

before treating dust. If one or both of the images are blurred (lower-quality optics, optical system out of focus, vibrations during the exposure), it may help to enhance constants p, q in the modified normalized cross-correlation function or to enhance parameter r_2 of the high-pass weight function.

4.3.5 Registration of similar images step by step

In this section, we will sum up the modified phase correlation for similar images. Let us denote f_1, f_2 the images to be registered. For computing the Fourier transform we will use a square $N \times N$ (N is an even number).

1. Multiplication of images f_1, f_2 by window functions, obtaining images f_{w1}, f_{w2}.
2. Centering images f_{w1}, f_{w2} in the square $N \times N$ pixels, obtaining images f_{c1}, f_{c2}.
3. Computing amplitude spectra A_1, A_2 of images f_{c1}, f_{c2}.
4. Computing logarithm of the amplitude spectra A_1, A_2 (precisely $\ln(1 + A_1), \ln(1 + A_1)$), obtaining functions A_{l1}, A_{l2}.
5. Transforming functions A_{l1}, A_{l2} to logarithmic-polar coordinates, obtaining functions A^{lp}_{l1}, A^{lp}_{l2}.
6. Multiplying functions A^{lp}_{l1}, A^{lp}_{l2} by rectangular window functions which diminish only at edges with $\rho = 0$ and $\rho = \frac{N}{2}$, obtaining functions A^{lp}_{w1}, A^{lp}_{w2}.
7. Computing the semi-normalized cross-power spectrum $Z^{p,q}_A$ of functions A^{lp}_{w1}, A^{lp}_{w2}. Constants p, q chosen close to zero.
8. Applying a low-pass high-pass weight function on Z_A, obtaining function Z_{wA}.
9. Computing the inverse discrete Fourier transform P_A of function Z_{wA}
10. Finding the coordinates of the global maximum of function P_A, i.e. finding the vector $(\ln \alpha, \theta)$.
11. Applying scale-change by α^{-1} and rotation by $-\theta$ to image f_{c2}, obtaining image f_3.
12. Computing the semi-normalized cross-power spectrum Z of images f_{c1}, f_3. Constants p, q chosen close to zero.
13. Applying a low-pass high-pass weight function on Z, obtaining function Z_w.
14. Computing the inverse discrete Fourier transform P of function Z_w.
15. Finding the coordinates of the global maximum of function P, i.e. finding the vector $(-x_0, -y_0)$.

The algorithm involves six discrete Fourier transforms and two inverse discrete Fourier transform computed by means of the Fast Fourier Transform algorithm. The number of Fourier transforms can be reduced by one if we keep the Fourier transform of image f_1 from step 3 to step 12. The discussion on the parameters of the algorithm is quite the same as for the shifted images, however, most of the parameters are used twice and they do not need to be the same in both cases.

4.3.6 Registration of motion-blurred images

It sometimes happens, not only in solar eclipse photography, that taken images are motion-blurred due to vibrations of the mount or motion of the observed objects. The latter happens for instance if a paralactic mount is not set up properly or a fixed mount is used for astronomical photography with too long exposure times. Both causes result in motion along a circle, but the curvature is usually negligible and the motion can be approximated by motion along a straight line. The discussion of the dynamics of this motion is more complicated. Vibrations of the mount can be taken as oscillation with attenuated sine course. If they are caused by the mirror or the shutter in the camera, the oscillation amplitude is the highest shortly after the exposure starts and decreases quickly. Another source of oscillations my be strong wind, which causes oscillations with the same frequency (eigenfrequency of the system), but unpredictable amplitude. The exposure time has a big impact on the blur of the images. Images with shorter exposure times are less likely blurred. Images with exposure time of a few seconds may be taken during several periods of the oscillation. Despite the lower quality of the images caused by motion blur, it is sometimes necessary to work with such images. Due to the complicated dynamics, it is impossible to use deblurring methods based on deconvolution.

Figure 4.10 illustrates the impact of motion blur of one of the images to be registered on the phase-correlation function. The global maximum in phase-correlation functions with higher values of λ is approximately at coordinates $(1, 1.73)$, which is the mean shift of the images used for creation of image f_2.

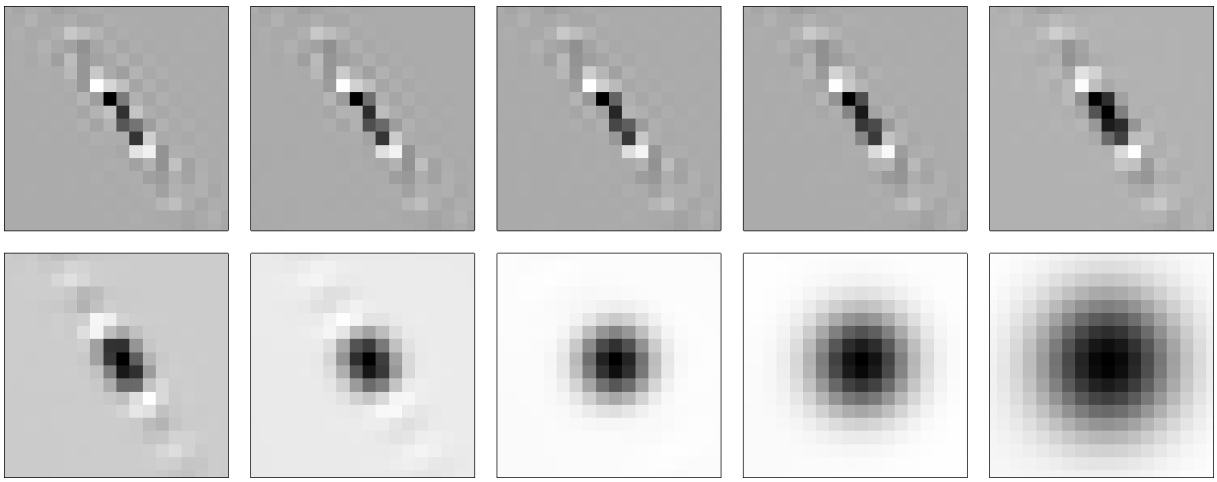


Figure 4.10: Details of peaks of the phase-correlation function of two images, one is a motion-blurred replica of the other one. Gaussian low-pass weight function was applied on the normalized cross-power spectrum. Parameter λ was set to $0, 1, 2, 4, 8, \ldots, 256$ (from top left to bottom right). The images are shown in negative, i.e. the darkest pixels are global maxima. Image f_1 is identical with the image used in Figure 4.8. To create image f_2, 17 replica of image f_1 were created with shift $\left(t\cos\frac{\pi}{3}, t\sin\frac{\pi}{3}\right), t = 0, 0.25, 0.5, \ldots, 4$, and they were averaged. A circular Gaussian window function and the T_8 filter were applied on both images (see Section 4.6 for details on the filter). The details show 18×18 pixels.

In images with no sharp contrasty structures (like Figure 4.12b on page 77), it is hard to say visually if the image is motion-blurred or not. The shape of the phase-correlation function peak can be a good tool for this assessment.

4.4 Subpixel precision registration

The modified semi-phase correlation function is represented by a matrix, which means that the searched coordinates (x_0, y_0) are integer-valued. Let (x_0, y_0) be the integer-valued estimate of the shift vector between images f_1 and f_2 computed as the coordinates of the global maximum of the semi-phase correlation function (with some modifications to the method described in the Section 4.3). We will compute the sub-pixel precision estimate of the shift vector, $(\bar{x}_0, \bar{y}_0)$. There are several methods how to estimate non-integer valued shifts.

Oversampling the images (described briefly in [2]) cannot enhance the registration precision much and leads to immense storage requirements and time-consuming computations. [2] describes fitting of a quadratic, Gaussian or modified sinc function through points $P(x_0, y_0)$, $P(x_0 + 1, y_0)$, $P(x_0 - 1, y_0)$, $P(x_0, y_0 + 1)$, $P(x_0, y_0 - 1)$, where P is the phase-correlation function, separately in each direction. These methods only use a small number of points around (x_0, y_0) and thus are quite sensitive to small variations in values of P in these points. On the other hand, the methods are based on the properties of the phase-correlation function of two replicas of the same image mutually shifted by a non-integer shift (Figure 4.8), which is a modified sinc function [43]. Therefore, they lead to very good results if the images to be registered are "nearly identical", e.g. video-rate images taken with a hand-held camera. As we can see from Figures 4.9 and 4.10, it is quite reasonable to take more points into consideration, especially in cases when the images are "more different". This is what the following method does.

A method described in [23] based on geometric moments turned up to be very robust. The sub-pixel precision estimate of the shift vector, $(\bar{x}_0, \bar{y}_0)$ is computed as

$$(\bar{x}_0, \bar{y}_0) = \left(\frac{M_{1,0}}{M_{0,0}}, \frac{M_{0,1}}{M_{0,0}} \right),$$

where $M_{k,l}$ is the geometric moment computed over a circle with center (x_0, y_0) and radius $\varepsilon \in \mathbb{R}^+$, i.e.

$$M_{k,l} = \sum_{x^2+y^2<\varepsilon} \sum x^k y^l P(x_0 + x, y_0 + y), \qquad k, l = 0, 1,$$

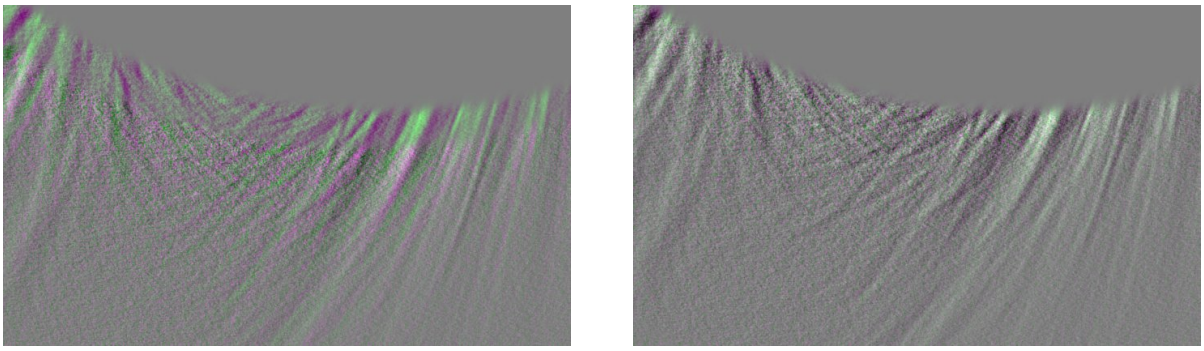
where $P(x, y)$ is the modified semi-phase correlation function from the last step of the sequence of steps described in Section 4.3.4. We may consider the point $(\bar{x}_0, \bar{y}_0)$ as the center of gravity of the peak and its neighborhood with radius ε. Usual values of ε range from 3 to 8. This parameter is another parameter which needs to be set manually when two image are to be registered.

The subpixel extension of the modified phase-correlation method describes estimating of a non-integer shift between images. It can also be used for enhancing the precision of the estimation of rotation and scale-change between similar images. The only difference is that the integer-valued vector (x_0, y_0) is replaced by a value of $(\ln \alpha, \theta)$ corresponding to integer indices in the matrix representing the modified phase-correlation function. The subpixel estimation gives us non-integer indices in the matrix, which correspond to a more precise value of $(\ln \alpha, \theta)$.

4.5 Checking the results

It is always very important to check if the computed registration parameters are correct. A simple method is to create a color composition of both images after one of them was transformed with the estimated parameters. There are several possibilities. One of them is to create an image with image f_1 in its red and blue color channels and transformed image f_2 in its green color channel. The human eye is the most sensitive in green [24], therefore both images are using about the same brightness range. Figure 4.11 is an example of images used for checking the registration results.

If both original images are identical up to a shift, the color image is neutrally gray. Images of the same "scene" taken at the same conditions with the same settings only show color noise. If the images are taken at different conditions, with different settings of the camera or the optical system, the image is not neutrally gray. Our knowledge about the content of the images will enable us to assess if they are registered correctly.



(a) (b)

Figure 4.11: Color images used for checking the registration results. Figure (a) shows a small part of images f_1, f_2 before registration, in Figure (b) image f_2 was shifted by the computed shift vector $(14.089, -40.955)$. The used images were taken by Constantinos Emmanoulidis in Akademgorodok, Novosibirsk, Russia on August 1, 2008 with digital camera Canon EOS 400D and Takahashi TSA102 814 mm lens.

4.6 Registration of solar corona images

Digital images of the solar corona taken during total solar eclipses are specific in many ways. The solar corona has very high contrast, which causes that it is impossible to capture it in a single image. It is necessary to take images with different exposure times to cover whole its brightness range. Images taken with very short exposure times are mostly underexposed with only a small part of the image correctly exposed. Images taken with longer exposure times have a part which is overexposed (saturated), a part which is correctly exposed, and a part with very low pixel values containing almost no information. Images taken with the longest exposure times have a large overexposed part and the rest is correctly exposed. We need to register images taken with different exposure

times and then compose all these images into one image with very high dynamic range. Specific methods used for solar corona image registration are described in [23].

Usually, there are not many structures visible in the original images. There is very high contrast, which may be eliminated by using a suitable high-pass weight function on the normalized cross-power spectrum. However, there are aspects that cause that it is better to replace the high-pass weight function by another procedure. The most contrasty structure in solar corona images from total solar eclipses is the edge of the Moon. The Moon moves in front of the Sun during the eclipse, which causes that registration of the images with respect to the Moon does not give the same results as registration with respect to the corona. If we do not remove the Moon edge from the images used for image registration, the images will be registered to the Moon and the corona will be registered incorrectly. Furthermore, a contrasty and sharp structure in the images is the edge of the saturated part of the image. Any non-linearity (edge of the saturated part, inhomogenous diffuse light) highly influences spatial frequencies in the radial direction and makes them unreliable for image registration.

4.6.1 Tangential high-pass filter T_ς

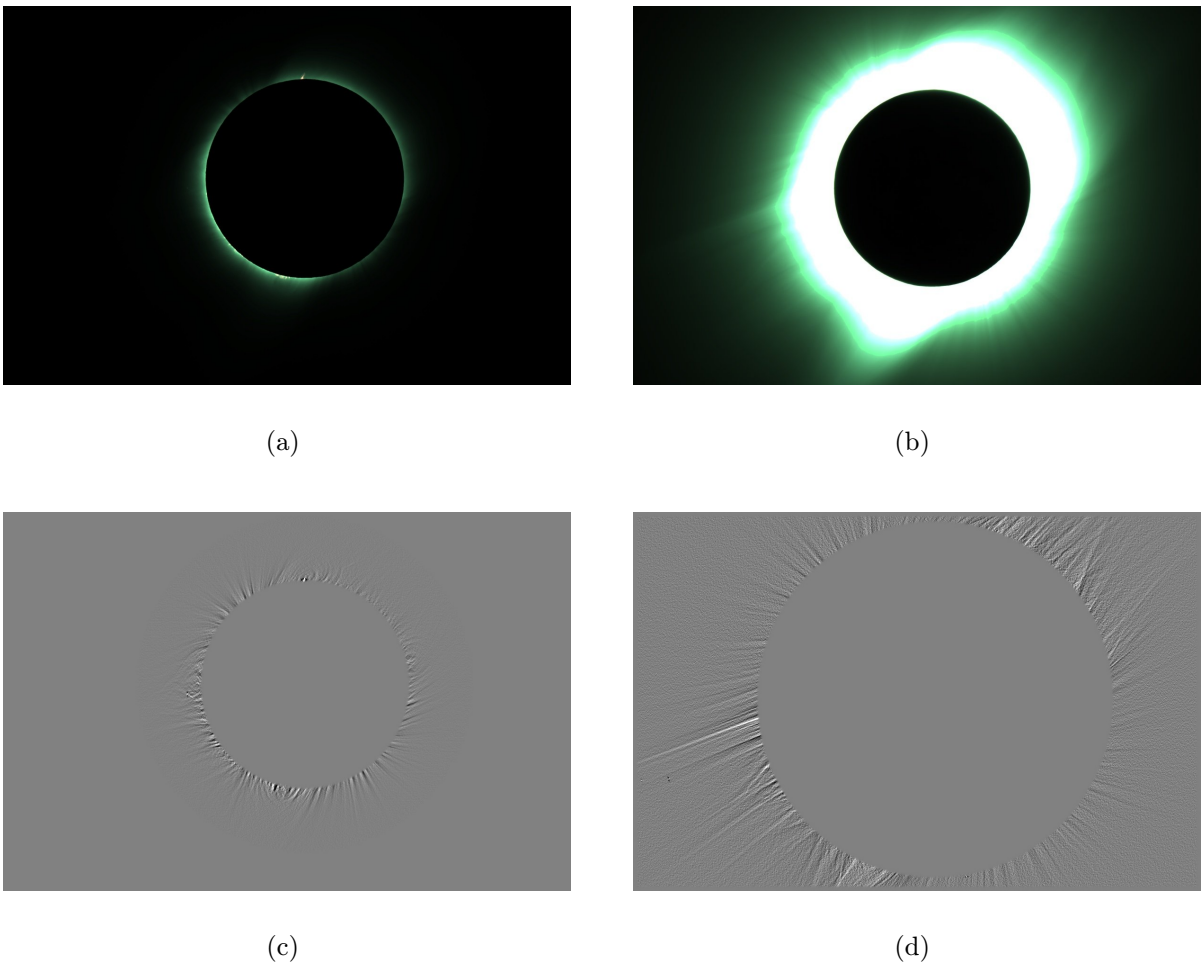
A solution to this problem is removing the information on spatial frequencies in radial direction and preserving the information on the spatial frequencies in tangential direction only. This cannot be done by applying a weight function on the spectrum. This must be done by a filter $T_\varsigma(f(x,y))$ applied on the images which is defined by the following formula

$$T_\varsigma(f(x,y)) = f(x,y) - \frac{1}{\varsigma\sqrt{2\pi}} \int_{l_1}^{l_2} f\left(x_c + \rho\cos\left(\varphi + \frac{l}{\rho}\right), y_c + \rho\sin\left(\varphi + \frac{l}{\rho}\right)\right) \mathrm{e}^{-\frac{l^2}{2\varsigma^2}}\,\mathrm{d}l, \tag{4.1}$$

where ρ, φ are the polar coordinates of point (x,y). The origin (x_c, y_c) of the polar coordinate system should be in the center of the Sun, which is not easy to be found. Therefore the center of the Moon is used instead as its approximation. The difference between the positions of the two centers is not big and it has no influence on the precision of image registration. The integral in formula (4.1) may be understood as an unsharp mask created by means of a one-dimensional low-pass filter with Gaussian kernel applied on the circle centered at the center of the Sun. Limits l_1, l_2 are usually set to $l = -3\varsigma, l_2 = 3\varsigma$. The difference between integration of the Gauss function from $-\infty$ to ∞ and from -3ς to 3ς is negligible since we work with 16-bit gray-scale images. The value of ς must be chosen according to the image quality. For sharp images, we choose e.g. $\varsigma = 8$, for less sharp images, we increase ς slightly. Since $f(x,y)$ is defined only for integer-valued (x,y), its values must be interpolated to compute the integral, the final value is then rounded for each pixel. Such a filter gives both positive and negative pixel values, therefore the image after filtration is stored in a signed type array, e.g. $\{-32\,768, \ldots, 32\,767\}$.

We must also apply a window function on the image. Due to the shape of the solar corona, circular window functions are usually more suitable. Computing a filtered image with a circular window function is much faster than computing a filtered image with a rectangular window function applied, because we do not need to compute the filtered image in a large part of the image if we use a circular window function. Nevertheless, if the exposure time of the image was long and a large part of the image is saturated, there is

only little space left in the annulus between the edge of the saturated part and the image edge. Then it is better to use a rectangular window function so that we use the image structures far from the center of the Sun. Moreover, it is necessary to remove the edge of the Moon and the edge of the saturated part. Therefore, the window function needs to be more like the low-pass high-pass weight function from Definition 4.19 (page 66) with $r_1 - \sigma_1$ at the edge of the saturated part if there is one. If there is no saturated part, then on the edge of the Moon. An example of original images of the solar corona (in linear brightness scale) and test images used for the estimation of scale-change, rotation and shift is in Fig. 4.12.



(a) (b)

(c) (d)

Figure 4.12: Example of solar corona images taken during a total solar eclipse. Images were taken by Constantinos Emmanoulidis in Akademgorodok, Novosibirsk, Russia on August 1, 2008 with digital camera Canon EOS 400D and Takahashi TSA102 814 mm lens. The exposure times were 1/60 s (image a) and 2 s (image b). Image (c) was created from image (a) using filter T_8 and a circular window function, image (d) was created from image (b) using filter T_8 and a rectangular window function. The exposure time of image (b) was so long that it was necessary to use a rectangular window function.

We use images computed using the described filtration and window function as images f_{w1}, f_{w2} in the estimation of the image rotation, rescale, and shift described step-by-step in Sections 4.3.4 and 4.3.5. The filtration is performed instead of the high-pass weight function, therefore no more high-pass weight functions are used.

Computing time

Computation of a test image for registration using the T_8 filter and a circular Gaussian window function takes 144 s for an $3\,472 \times 2\,314$ pixel image with the diameter of the Sun 986 pixels on a computer with Intel Pentium D 3.2 GHz CPU and 800 MHz memory. To compute the test image with a rectangular window function, it takes 206 s.

4.6.2 Methods and parameters used for solar corona images

Miloslav Druckmüller developed the first version of the computer program for image registration based on phase correlation, PhaseCorr, based on the article [45]. The current version is PhaseCorr 6.0. The control window of the PhaseCorr program is in Figure 4.13. I have been using the program for several years and gained some skills about the shape of the peak in the phase-correlation function and the parameters which influence it. In this section, I would like to mention which methods described in the previous sections we use and how we do that.

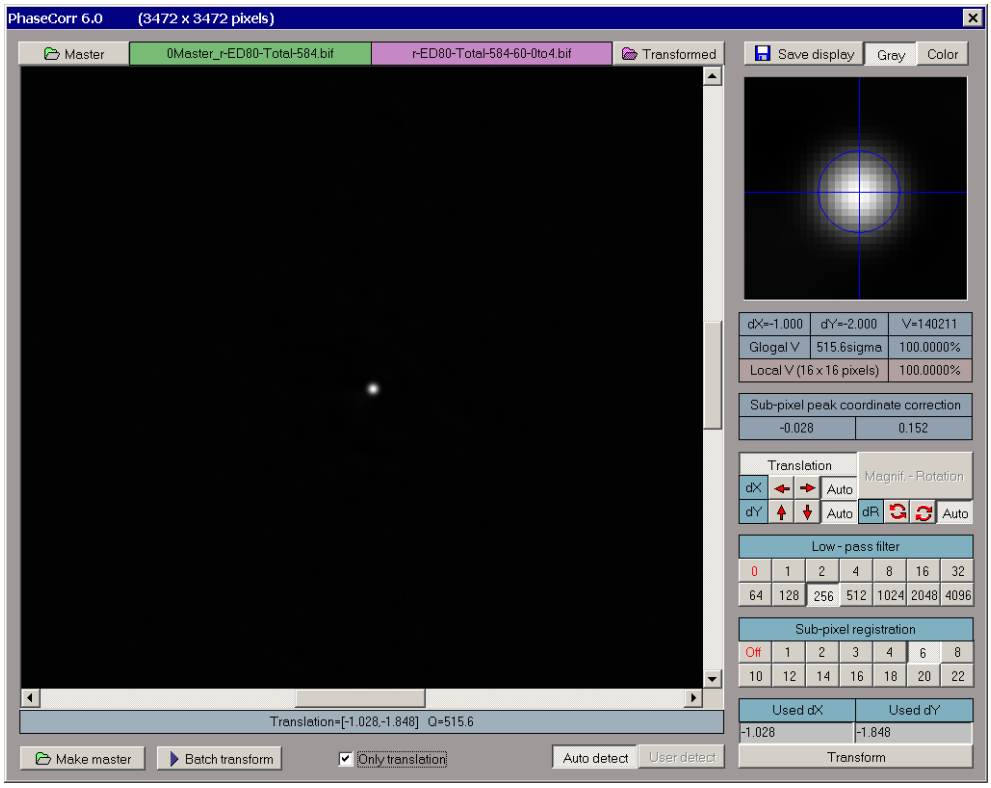


Figure 4.13: The control window of PhaseCorr 6.0.

Let us suppose that we have two images f_1, f_2 of the solar corona taken during a total solar eclipse and we would like to register them. They must be taken with the same or similar exposure times (up to about two stops depending on their quality) or if they are

taken with different lenses, their exposure times must be similar in the sense of how much the corona is exposed. To register images with very different exposure times, we need to have some images with exposure times between f_1 and f_2 which are registered to one of the images (e.g. the one with shorter exposure time, f_1). Then we average all images we already have registered, transformed by the computed similarity transformation. Some of these images have exposure times between f_1 and f_2. This average is then used instead of image f_1 for the registration.

First we create so called test images used only for the similarity parameter estimation. This is done in the Corona software. I created the first version of this program in 2002 and since that time, Miloslav Druckmüller has developed it into a complicated software for coronal structure enhancement. In the Corona software, we set the approximate position of the Moon in the images and its radius. Then the program computes the test images applying a circular or rectangular window function and the T_ς filter. For high quality images, we use $\varsigma = 8$, for good images, $\varsigma = 9$ and for lower quality images, we use $\varsigma = 10$.

These images are used for registration in the PhaseCorr program. It uses phase correlation, i.e. parameters $p = q = 0$ in the formula for the semi-phase correlation function. It only performs a check if the product of amplitude spectra of both images is non-zero in each pixel. If it is zero, it divides by one instead of zero.

As the low-pass weight function, it uses the Gaussian low-pass weight function with $\lambda = 0, 1, 2, 4, \ldots, 4096$. Commonly used values are 8 to 256. Higher values are used only for poor quality images and images with bad correspondence. The recent version also enables to use subpixel precision registration using the method based on geometrical moments described in Section 4.4 (page 74).

When the rotation angle θ and scale factor α are computed, the transformation is performed with the test image only. The whole similarity transformation is performed after the shift vector is estimated, which keeps as large part of the original image as possible.

Computing time

For computing the discrete Fourier transform, PhaseCorr uses the FFTW2dll library written by Jindřich Nový, which is sometimes even faster than the Fast Fourier Transform. Registering shifted images, the PhaseCorr program performs steps 3 to 6 from the algorithm in Section 4.3.4. On a computer with Intel Pentium D 3.2 GHz CPU, 800 MHz memory, the computation takes 21 s for $N = 3\,472$. To compute one Fourier transform, it takes 6.2 s. Registering similar images, the program performs steps 3 to 10 from the algorithm in Section 4.3.5. After the user chooses parameters λ and ε, the test image for image f_2 is rotated and scaled by the computed parameters and the user proceeds to registration of images which are shifted only. Steps 3 to 10 take 44 s for the image and computer described above.

4.6.3 Registration of images at different wavelengths

In order to comprehend better the processes in the solar corona, it is very useful to study the radiation of ions or atoms at specific spectral lines. Most of the radiation of the corona is the photospheric light scattered by free electrons in the corona. Its intensity depends on the wavelength by the Planck law [29] for temperature 5 785 K (the average temperature of the photosphere) [4]. For this purpose, it is necessary to take photographs at the

wavelength of the spectral line through a narrow-band filter. Apart from these so called *on-line images*, it is necessary to take so called *off-line images*, which are images taken through a narrow-band filter with the same band-width and the transmission shifted about a nanometer from the spectral line. The difference between the photospheric radiation at these two wavelengths can be considered as negligible. First, the on-line and off-line images are registered separately, then each images of exposure time is calibrated by flat-field images and the off-line images for each exposure time are subtracted from the on-line images. Thus we obtain images showing radiation of the studied ions or atoms only. These images are then composed into one image with high dynamic range and coronal structures are enhanced.

If we want to register images showing the radiation of different ions or atoms at different wavelengths, there may be few similar structures in the images. However, we can use the off-line images used for creation of both images to estimate the transformation parameters. Both sets of off-line images contain about the same structures since they show the photospheric light scattered by free electrons in the corona. Thus we use very similar off-line images to estimate the parameters of the similarity between very dissimilar images. Examples of images showing the radiation of specific iron ions (on-line minus off-line images) and their off-line images are in Figure 4.14. The off-line images 4.14b,d contain similar structures and hence can be used for registration of images 4.14a,c.

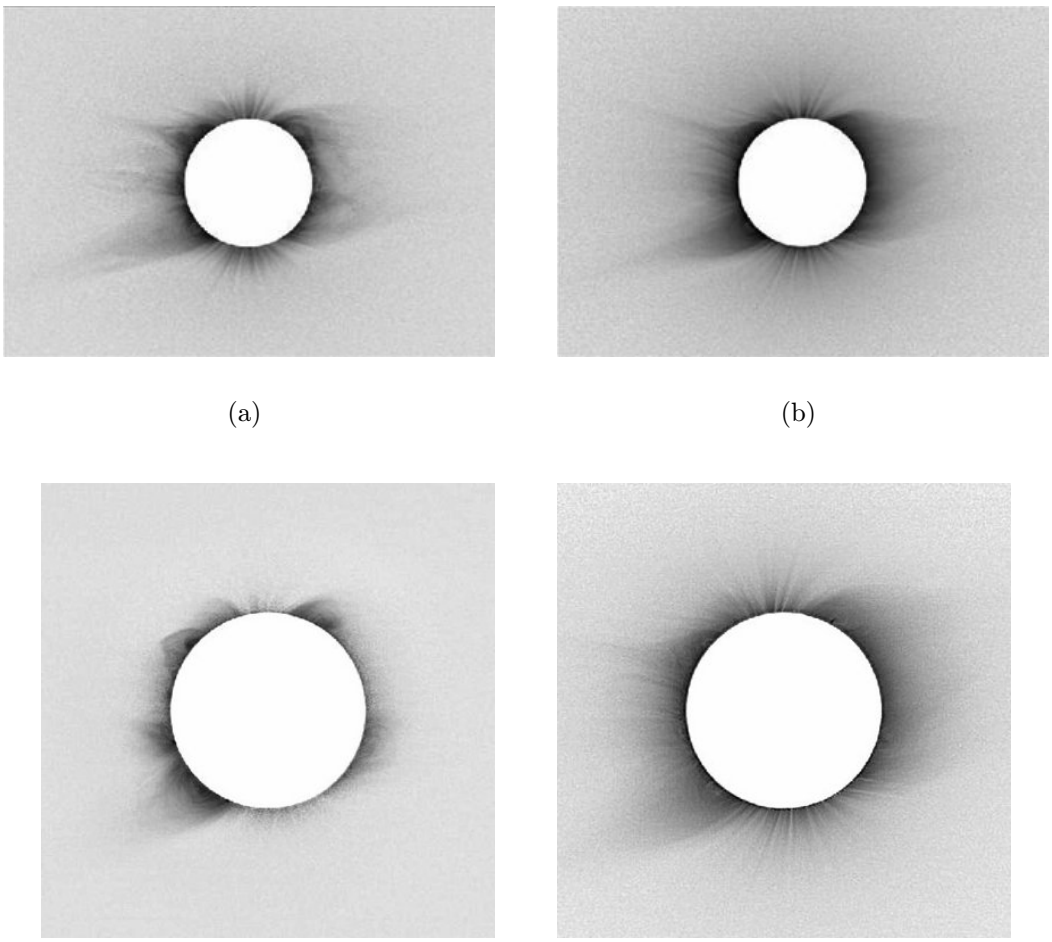
4.7 Testing registration precision

4.7.1 Testing on simulated data

I generated 30 realizations (X_k, Y_k) of the random vector (x_0, y_0) standing for the shift vector between two images. Both x_0 and y_0 were uniformly distributed over $\langle -200, 200 \rangle$. The generated shift vectors were rounded to three decimals obtaining vectors (x_k, y_k). Image f_1 was a high quality image used in Figure 4.8 on page 69. The width of the image was $N = 3\,472$, which was also the size of the matrix used for the discrete Fourier transform. Image f_2 was shifted with respect to image f_1 by vector (x_k, y_k), $k = 1, 2, \ldots, 30$. The subpixel shift was performed by means of bilinear interpolation. A circular Guassian window function was applied and the image used for shift estimation was filtered by the T_8 filter (see Section 4.6.1, page 76). I tried to find the used shifts using the subpixel extension of the phase correlation method. Knowing what the correct shift was, I tried to set parameters r_2 for the low-pass weight function and ε for subpixel shift estimation so that the estimated shift is as close to (x_k, y_k) as possible. Using $r_2 = 32$ and $\varepsilon = 8$, I succeeded to find precisely the shift vector (x_k, y_k) for all $k = 1, 2, \ldots, 30$.

Then I did the same test with rotation and scale-change. The generated vectors (α_k, θ_k), $k = 1, 2, \ldots, 30$ were used for the creation of image f_2 scaled by factor α_k and rotated by θ_k degrees with respect to image f_1 using bilinear interpolation. $\alpha_k \in \langle 0.9, 1.1 \rangle$ with four decimals, $\theta_k \in \langle -10, 10 \rangle$ with three decimals. It was necessary to vary the values of r_2 in the range of $16, 32$ and 64 and the values of ε in the range $4, 6, 8, 10$ to obtain the best estimates. In 6 of 30 cases, it was impossible to find α_k or θ_k exactly. However, the error was only on the fourth decimal for α_k (one case) and on the third decimal for θ_k (five cases).

Testing on simulated data showed that the phase correlation method with its modifications and the subpixel precision extension are very precise tools for estimating parameters



(a) (b)

(c) (d)

Figure 4.14: Images taken during the total solar eclipse on August 1, 2008 with narrow-band filters. Figure (a) shows the radiation of Fe XI ion, Figure (c) the radiation of Fe XIV ion, they are on-line images with off-line images subtracted. Figures (b) and (d) are the off-line images for images (a) and (c). All images are compositions of many images taken with different exposure times, registered by modified phase correlation, composed in a single image with very high dynamic range. Finally, structure enhancement was applied on the images. The images are shown in negative. They are taken from [30].

of similarity between images. In practice, the precision of registration depends on quality of the used images. The precision of the method is much higher than the precision which is possible due to other factors influencing the image, e.g. seeing, vibrations of the mount, image blur caused by movement of the objects which are captured etc.

Testing on images with simulated additive noise

I tested the registration precision on images with added additive noise too. I took the image used for the testing in the previous section and shifted it by $(16, 17)$. Then I reduced the size of both the original image and the shifted image three times. I cropped a square 752×752 pixels from both images and added additive noise with normal distribution

to both the images. I chose this procedure so that there is no systematical error in the registration caused by the organization of the Bayer mask before the sensors. I cropped both images after the resize so that there is no missing part filled with zeroes in the shifted image. This enabled me to use the same rectangular Gaussian window function on both the images.

Then I used the modified phase-correlation method to estimate the shift between the images. The following table shows the results. The values of ω are standard deviations of the normal distribution used for the additive noise simulation. λ and ε are the values of the parameter of the low-pass function (Definition 4.20 on page 68) and the radius used for subpixel correction (Section 4.4 on page 74) which gave the estimated shift closest to $\left(\frac{16}{3}, \frac{17}{3}\right) \doteq (5.333, 5.667)$. The last column shows the error of such estimation, i.e. the Euclidean norm of the difference of the correct and estimated shift vector. Parameter ς was set to 8 and the registration precision was insensitive to this parameter. $\omega = 32$ is about the lowest standard deviation of additive noise which Canon EOS 5D digital camera has at ISO 100. $\omega = 512$ is about the highest standard deviation of additive noise which the camera has at ISO 3 200. The additive noise standard deviation depends on temperature.

ω	λ	ε	error
32	16	4	0.023
64	32	6	0.053
128	32	6	0.084
256	64	8	0.027
512	64	6	0.942

The table shows that the more significant the additive noise is, the more we should enhance λ to filter out information on high spatial frequencies deteriorated by the additive noise. The influence of additive noise on the registration precision is almost negligible provided that images were not taken with some of the highest ISO setting the camera offers.

4.7.2 Testing on real data

It is problematic to test the precision of registration on real solar eclipse data, because we would need another method with comparable or higher precision which would be able to give us parameters of the geometrical transformation of images. [23] uses a method based on measurement of the lunar centroid. For each image, we need to know precisely the time when the image was taken. We can assume that the mutual shift of the images to be registered is the shift of the lunar centroid corrected by the shift of the Moon with respect to the Sun in the images computed from the time when the images were taken, used optics and parameters of the camera sensor. Only images with short exposures can be used, because overexposed parts of the image at the lunar edge would lead to imprecise lunar centroid measurement.

The testing in [23] was performed on images taken on August 1, 2008 in Mongolia by Miloslav Druckmüller, Peter Aniol and Vojtech Rušin with Canon EOS 5D digital camera and 1250 mm lens. The following parameters of the modified phase correlation were used: $N = 4\,096$, $p = q = 0.01\,\% \max\{A_{f_1}, A_{f_2}\}$, $r_2 = 0.1N$, $\varsigma = 8$, $\varepsilon = 3$. Since paralactic

mount was used, the mutual rotation of images was considered to be zero. The lens was not refocused between the images, therefore no scale-change occurred. Registration by means of modified phase-correlation and using the lunar centroid do not give differences higher than 0.40 pixel on the used images.

The data set which the testing in [23] was performed on is quite unique. It is hard to repeat such test with another data set, because the requirements on the images are very strong. First, we need to know the times when the images were taken with higher precision than one second (which is the highest precision available from the EXIF data of the files). Therefore, we need images taken using a computer program which can guarantee the precision. We use the Multican application for Linux created by Jindřich Nový. Secondly, we need images which are shifted with respect to each other, meaning that the paralactic mount should not be set up in the best possible way. Otherwise, there is nothing to measure by means of the phase-correlation based method. (Un)fortunately, the mount used in 2009 on Enewetak was so good and so well set up that there is no measurable shift among images taken tens of seconds one after another. Last but not least, only images taken with a suitable focal length must be used. The longer focal length we use, the more the Moon moves and the longer shifts can be measured. Finally, there are no better or comparable good data as those used in [23] which can be used for the testing.

Chapter 5

Conclusion

Throughout the thesis, I gave proves to several theorems which I have not found proves to in literature – Theorems 2.45, 2.49c,d, 3.14, 3.15, and 3.24.

Gathering theorems from various sources, I realized that some theorems are stated in a careless manner. Therefore, I added assumptions to Theorems 2.48, 2.49, 2.54, 2.55, 2.61, and 2.67.

I defined the notions of $\mathcal{L}(\mathbb{R})$, $\mathcal{L}(\mathbb{R}^2)$ spaces, semi-normalized cross-power spectrum, cross-correlation function, phase-correlation function and semi-phase correlation function, periodization of a function $\{0, 1, \ldots, N-1\}^2 \to \mathbb{C}, N \in \mathbb{N}$ and its Fourier spectrum, discrete Fourier transform and inverse discrete Fourier transform of periodized functions. I introduced functions $\zeta_{(1),p}(t)$ and $\zeta_{(2)}(s,t)$.

I stated and proved some new minor results in the Fourier analysis and phase-correlation theory, namely Theorems 2.38, 2.46, Corollary 2.57, Theorems 2.59, 2.64, 2.66, 3.17, 3.18, Corollary 3.21, Theorems 3.27, 3.29, and 3.31.

I described the method for registration of scaled, rotated and shifted images based on phase-correlation with its modifications which make it possible to use the method to register even quite dissimilar images. The images may have high dynamic range and faint structures, they may contain additive noise, impulse noise, faulty pixels, spots caused by dust particles. I described the practical aspects of image registration which cannot be found in literature in details. Namely, I described the need of removing image edges and suggested window functions for this purpose. The method for registration of rotated and scaled images is only described in [45] and I gave more rigorous mathematical derivations of the algorithms. I showed the reasons for applying weight functions on the Fourier spectra and suggested additive constants in the denominator of the phase-correlation formula in image with some kinds of modulation transfer function.

I described the extension of this method to subpixel precision and tested the precision. The method is so precise that other factors influencing image quality, like seeing, vibrations of the camera and motions of the observed object have much bigger effect. The method proved to be very robust and it can register images of the solar corona taken during total solar eclipses which are impossible to be registered manually by control points or other registration techniques.

The precise registration of images is very important. If the registered images are then composed in a single image, imprecise registration causes blur of the resulting image. Image registration with error smaller than one pixel is a necessary condition for recovering the faint structures in the visible solar corona in a way showed in Figure 5.1. Such images enable us to study the strong solar magnetic field visualized by the photospheric light

scattered on free electrons in the visible light in high resolution. Special techniques can be used to visualize different types of ions which map different temperatures. Without registration of there two types of images, it is impossible to use to full advantage to this data. Cosmic probes also study the solar corona. It is either using a coronagraph (the Sun is obscured by a disc with higher diameter than the solar disc) in the visible light or in the ultra-violet part of the spectrum. Coronagraphs cannot show the innermost part of the corona, where the most fascinating processes take place. Therefore, the solar eclipses are unique opportunities to study the inner corona in visible light.

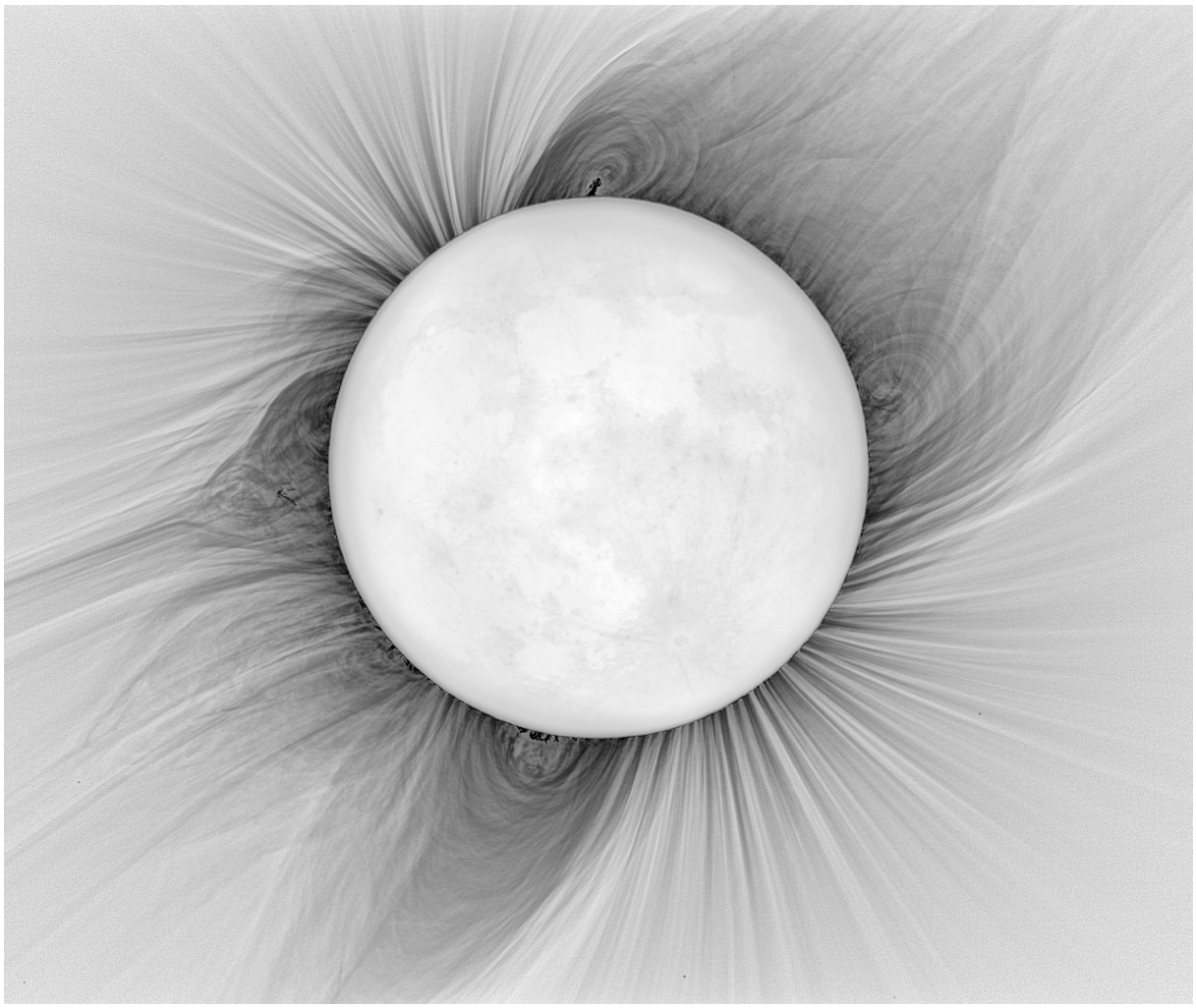


Figure 5.1: An image showing the inner corona during the total solar eclipse on August 1, 2008 (in negative). The images used for the creation of this image were taken by Constantinos Emmanoulidis in Akademgorodok, Novosibirsk, Russia with Takahashi TSA102 814 mm lens and Canon EOS 400D digital camera. 37 images with exposure times 1/250 s to 2 s were used. Eclipse images were calibrated by means of dark frames and flat-fields, registered by means of the modified phase-correlation method described in this thesis, composed by means of LDIC 5.0 software and processed using Corona 4.1 in order to visualize coronal structures. Final processing was done using ACC 6.1 software. Image processing by Hana Druckmüllerová and Miloslav Druckmüller [19]. The image is one of the best images showing the inner corona of the Sun.

Bibliography

[1] APICELLA, A. – KIPPENHAN, J. S. – NAGEL, J. H. Fast multi-modality image matching. *Medical Imaging III: Image Processing*, 1989, vol. 1092, pp. 252-263. Available from <http://elib.uni-stuttgart.de/opus/volltexte/2010/5192/pdf/nag10.pdf>

[2] ARGYRIOU, V. – VLACHOS, T. *A study of sub-pixel motion estimation using phase correlation* [online]. [cit. April 11, 2010]. <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.96.4780&rep=rep1&type=pdf>.

[3] ASADZADEH, M. *Lecture Notes in Fourier Analysis* [online]. Last modified September 1, 2008 [cit April 16, 2010]. <http://www.math.chalmers.se/~mohammad/teaching/Fourier/LectureNotes_A3/draft_1.pdf>

[4] ASCHWANDEN, M. J. *Physics of the Solar Corona: An Introduction with Problems and Solutions*. 2nd edition. Berlin (Germany): Springer-Verlag and Chichester (UK): Praxis Publishing Ltd, 2006. 892 p. ISBN 3-540-30765-6.

[5] ATKINS, B. *Mirror Lenses – how good? Tamron 500/8 SP vs Canon 500/4.5L* [online]. Last modified May 24, 2008, [cit. February 2, 2010]. <http://www.bobatkins.com/photography/tutorials/mirror.html>.

[6] BARTALOS, T. *Registrace obrazu se subpixelovou přesností*. Olomouc: Univerzita Palackého. Přírodovědecká fakulta. Katedra informatiky, 2007. 82 p. Master thesis supervisor Dobeš, M.

[7] BEERENDS, R. J., et al. *Fourier and Laplace Transforms.* Translated from Dutch by Beerends, R. J. Cambridge (New York): Cambridge University Press, 2003. 447 p. ISBN 0521806895.

[8] BEZVODA, V., et al. *Dvojrozměrná diskrétní Fourierova transformace a její použití – I.: Teorie a obecné užití.* 1st edition. Praha: Státní pedagogické nakladatelství, n. p., 1988. 181 p. ISBN 17-135-88.

[9] BORELL, C. *Lecture Notes in Measure Theory* [online]. c2006, last revision August 25, 2009 [cit. November 17, 2009]. <http://www.math.chalmers.se/Math/Grundutb/GU/MMA110/A09/MeasureTheory.pdf>.

[10] BORELL, C. *Integration Theory* (lecture). Gothenburg (Sweden): Chalmers University of Technology, Mathematical Sciences, September 22, 2009.

[11] BORELL, C. *Integration Theory* (lecture). Gothenburg (Sweden): Chalmers University of Technology, Mathematical Sciences, October 6, 2009.

[12] BRACEWELL, R. N. et al. Affine Theorem for Two-Dimensional Fourier transform. *Electronics Letters*, 1993, vol. 29, no. 3, pp. 304. ISSN 0013-5194.

[13] BROWN, L. G. A Survey of Image Registration Techniques. *ACM Computing Surveys*, 1992, vol. 24, no. 4, pp. 325–376. ISSN 0360-0300.

[14] BROWN, M. *Bayer mask CCD* [online]. Last modified July 2, 1999 [cit. April 16, 2010]. <http://www.nezumi.demon.co.uk/photo/bayer/bayer.htm>.

[15] ČÍŽEK, V. *Diskrétní Fourierova transformace a její použití.* 1st edition. Praha: SNTL – Nakladatelství technické literatury, n. p., 1981. 160 p. Matematický seminář SNTL. ISBN 04-019-81.

[16] DE CASTRO, E. – MORANDI, C. Registration of Translated and Rotated Images Using Finite Fourier Transforms. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1987, vol. PAMI-9, no. 5, pp. 700–703. ISSN 0162-8828.

[17] DĚMIDOVIČ, B. P. *Sbírka úloh a cvičení z matematické analýzy.* Translated from Russian by Rozložník, M., Tůma, M. 1st edition. Havlíčkův Brod: Fragment, 2003. 460 p. ISBN 80-7200-587-1.

[18] DRUCKMÜLLER, M. *Analýza v komplexním oboru.* 1st edition. Brno: Fakulta strojní, Vysoké učení technické v Brně, 1986. 96 p.

[19] DRUCKMÜLLER, M. *Inner corona* [online]. c2008, last updated March 19, 2009 [cit. April 16, 2010]. <http://www.zam.fme.vutbr.cz/~Eclipse/Ecl2008r/Tse2008_tsa102_1/0-info.htm>.

[20] DRUCKMÜLLER, M. *Numerické metody analýzy obrazů* (lecture). Brno: Brno University of Technology, Faculty of Mechanical Engineering, March 10, 2009.

[21] DRUCKMÜLLER, M. *Numerické metody analýzy obrazů* (lecture). Brno: Brno University of Technology, Faculty of Mechanical Engineering, March 17, 2009.

[22] DRUCKMÜLLER, M. *Numerické metody analýzy obrazů* (lecture). Brno: Brno University of Technology, Faculty of Mechanical Engineering, April 7, 2009.

[23] DRUCKMÜLLER, M. Phase correlation method for the alignment of total solar eclipse images. *The Astrophysical Journal*, 2009, vol. 706, no. 2, pp. 1605–1608. ISSN 0004-637X.

[24] DRUCKMÜLLEROVÁ, H. *A method for the visualization of high phase gradients in a microscopic image.* Brno: Brno University of Technology, Faculty of Mechanical Engineering, 2008. 31 p. Bachelor thesis supervisor Chmelík, R.

[25] DRUCKMÜLLEROVÁ, Z. *Korekce pozadí zobrazení fáze v digitálním holografickém mikroskopu užitím konvoluce.* Brno: Brno University of Technology, Faculty of Mechanical Engineering, 2010. 17 p. Bachelor thesis supervisor Uhlířová, H.

[26] FOLLAND, G. B. *Fourier Analysis and Its Applications.* Second Edition. Providence (Rhode Island, U.S.A.): The American Mathematical Society, 2009. 433 p. The Sally series, Pure and Applied Mathematics, Undergraduate Texts. ISBN 978-0-8218-4790-9.

[27] FOLLAND, G. B. *Real Analysis: Modern Techniques and Their Applications.* Second Edition. New York (U.S.A.): John Wiley & Sons, Inc., 1999. 386 p. Pure and Applied Mathematics, A Wiley-Interscience Series of Texts, Monographs, and Tracts. ISBN 0-471-31716-0.

[28] FOVEON. *Foveon* [online]. [cit. April 30, 2010]. <www.foveon.com>

[29] GOLUB, L. – PASACHOFF, J. M. *The Solar Corona.* Second edition. New York: Cambridge University Press, 2010. 390 p. ISBN 978-0-521-88201-9.

[30] HABBAL, S. R., et al. Mapping the distribution of electron temperature and Fe charge states in the corona with total solar eclipse observations. *The Astrophysical Journal*, 2009, vol. 708, no. 2, pp. 1650–1662. ISSN 0004-637X.

[31] HARRIS, F. J. On the Use of Windows for Harmonig Analysis with the Discrete Fourier Transform. *Proceedings of the IEEE*, January 1978, vol. 66, no. 1. ISSN 0018-9219.

[32] HOWELL, K. B. *The Transforms and Applications Handbook, Chapter 2: Fourier Transforms* [online]. c2000 [cit. November 4, 2009]. <http://www.mathnetbase.com/books/672/ch02.pdf>.

[33] KOLMOGOROV, A. N. – FOMIN, S. V. *Základy teorie funkcí a funkcionální analýzy.* From Russian Колмогоров, А. Н. – Фомин, С. В.: Элементы теории функций и функционального анализа, published by Наука, Moscow 1972, translated by Doležal, V., Tichý Z. 1st edition. Praha: SNTL – Nakladatelství technické literatury, n. p., 1975. 584 p. Teoretická knižnice inženýra. ISBN 04-01575.

[34] KOMRSKA, J. *Fourierovké metody v teorii difrakce a strukturní analýze.* 1st edition. Brno: Vysoké učení technické v Brně, Nakladatelství VUTIUM, 2001. 222 p. ISBN 80-214-2011-1.

[35] LIKEŠ, J. – MACHEK, J. *Počet pravděpodobnosti.* 1st edition. Praha: SNTL – Nakladatelství technické literatury, n. p., 1982. 160 p. ISBN 04-010-82.

[36] MAES, F. et al. Multimodality Image Registration by Maximization of Mutual Information. *IEEE Transactions on medical imaging*, 1997, vol. 16, no. 2, pp. 187–189. ISSN 0278-0062.

[37] NOVÝ, J. *Digital filtering and compression in image processing and volume rendering.* Brno: Brno University of Technology, Faculty of Mechanical Engineering, 2005. 90 p. PhD thesis supervisor Druckmüller, M.

[38] REKTORYS, K. et al. *Přehled užité matematiky.* 1st edition. Praha: SNTL – Nakladatelství technické literatury, n. p., 1963. 1136 p. ISBN 04-018-63.

[39] ROSENGREN, H. *Laplace Transform* [online]. c2008, last revision September 2, 2009 [cit. December 1, 2009]. <http://www.math.chalmers.se/Math/Grundutb/GU/MMG710/H09/laplace.pdf>.

[40] ROSENGREN, H. *Fourier Analysis* (lecture). Gothenburg (Sweden): Chalmers University of Technology, Mathematical Sciences, October 5, 2009.

[41] ROSENGREN, H. *Fourier Analysis* (lecture). Gothenburg (Sweden): Chalmers University of Technology, Mathematical Sciences, October 12, 2009.

[42] SCHWARTZ, L. *Matematické medody ve fyzice*. From French Méthodes mathématiques pour les sciences physiques, Paris 1965, translated by Brabec, J. 1st edition. Praha: SNTL – Nakladatelství technické literatury, n.p., 1972. 360 p. Teoretická knižnice inženýra. ISBN 04-018-72

[43] SHEKARFOROUSH, H. – BERTHOD, M. – ZERUBIA, J. Subpixel Image Registration by Estimating the Polyphase Decomposition of the Cross Power Spectrum. *Rapport de recherche*, November 1997, no. 2707. pp. 1-24. ISSN 0249-6399.

[44] SNEDDON, I. N. *Fourier Transforms*. 1st edition. New York: McGraw-Hill Books Company, Inc., 1951.

[45] SRINIVASA REDDY, B. – CHATTERJI, B. N. An FFT-Based Technique for Translation, Rotation, and Scale-Invariant Image Registration. *IEEE Transactions on Image Processing*, August 1996, vol. 5, no. 8, pp. 1266-1271. ISSN 1057-7149.

[46] STEIN, E. M. – WEISS, G. *Introduction to Fourier Analysis on Euclidean Spaces.* 1st edition. Princeton (New Jersey, U.S.A.): Princeton University Press, 1971. 297 p. ISBN 0-691-08078-X.

[47] STONE, H. S. et all. A Fast Direct Fourier-Based Algorithm for Subpixel Registration of Images. *IEEE Transactions on Geoscience and Remote Sensing*. October 2001, vol. 39, no. 10, pp. 2235-2243. ISSN 0196–2892.

[48] VERSTRAETE, J. *Lecture 16* [online]. cit. 17 January, 2010. <http://www.math.ucsd.edu/~jverstra/20e-lecture16.pdf>.

[49] ZITOVÁ, B. – FLUSSER, J. Image registration methods: a survey. *Elsevier Image and Vision Computing*, 2003, vol. 21, pp. 977—1000. ISSN 0262-8856.

[50] *Comparaison des Canon 40D, 50D, 5D et 5D Mark II* [online]. [cit. 5 February, 2010]. <http://www.astrosurf.com/buil/50d/test.htm>.

Used symbols

$\mathbb{N}$	the set of numbers $1, 2, 3, \ldots$
$\mathbb{N}_0$	the set of numbers $0, 1, 2, \ldots$
$\mathbb{Z}$	the set of integers
$\mathbb{R}^+$	the set of all positive real numbers, i.e. $(0, \infty)$
$\mathbb{R}_0^+$	the set of all non-negative real numbers, i.e. $\langle 0, \infty)$
$\mathcal{L}(\mathbb{R})$	space of all function $\mathbb{R} \to \mathbb{C}$ with finite integral of $\lvert f \rvert$, see Definition 2.4
$\mathcal{L}(\mathbb{R}^2)$	space of all function $\mathbb{R}^2 \to \mathbb{C}$ with finite integral of $\lvert f \rvert$, see Definition 2.8
C^1	class of continuous functions with continuous derivatives, see Definition 2.30
a^*	the complex conjugate of $a \in \mathbb{C}$
$\lfloor a \rfloor$	the integral part of real number a, see Definition 2.40
$\oplus$	addition modulo 2π, see Definition 2.41
$\chi_A(x)$	the characteristic function of set A, see Definition 2.10
$\mathcal{F}$	the Fourier transform, see Definitions 2.21, 2.34
$\mathcal{F}^{-1}$	the inverse Fourier transform, see Definitions 2.23, 2.35
$\mathcal{D}$	the discrete Fourier transform, see Definitions 3.1, 3.8
$\mathcal{D}^{-1}$	the inverse discrete Fourier transform, see Definitions 3.2, 3.9
$f(x,y), f_1(x,y), f_2(x,y)$	functions from $\mathcal{L}(\mathbb{R}^2)$ or functions $\{0, 1, \ldots, N-1\} \to \mathbb{C}, N \in \mathbb{N}$
N	size of the domain functions $f(x,y), f_1(x,y), f_2(x,y)$, which are defined on $\{0, 1, \ldots, N-1\}, N \in \mathbb{N}$, N is supposed to be an even number
(P.V.)	the Cauchy principal value, see Definition 2.28
$F(\xi,\eta), F_1(\xi,\eta), F_2(\xi,\eta)$	the Fourier spectra of functions $f(x,y), f_1(x,y), f_2(x,y)$
$A(\xi,\eta), A_1(\xi,\eta), A_2(\xi,\eta)$	the amplitude spectra of functions $f(x,y), f_1(x,y), f_2(x,y)$
$\Phi(\xi,\eta), \Phi_1(\xi,\eta), \Phi_2(\xi,\eta)$	the phase spectra of functions $f(x,y), f_1(x,y), f_2(x,y)$
$C_{f_1,f_2}(\xi,\eta)$	the cross-power spectrum of functions f_1, f_2, see Definitions 2.62, 3.25
$Z_{f_1,f_2}(\xi,\eta)$	the normalized cross-power spectrum of functions f_1, f_2, see Definitions 2.62, 3.25
$Z^{p,q}_{f_1,f_2}(\xi,\eta)$	the semi-normalized cross-power spectrum of functions f_1, f_2, see Definitions 2.62, 3.25
$Q_{f_1,f_2}(\xi,\eta)$	the cross-correlation function of functions f_1, f_2, see Definitions 2.63, 3.26

$P_{f_1,f_2}(\xi,\eta)$	the phase-correlation function of functions f_1, f_2, see Definitions 2.63, 3.26
$P^{p,q}_{f_1,f_2}(\xi,\eta)$	the semi-phase correlation funcion of functions f_1, f_2, see Definitions 2.63, 3.26
p, q	parameters of the semi-normalized cross-power spectrum and the semi-phase correlation function, see Definitions 2.62, 2.63, 3.26, 3.26
(x_0, y_0)	the shift vector between functions/images f_1, f_2, see e.g. the Shift Theorem 2.42
α	the scale-change factor between functions/images f_1, f_2, see e.g. the Scale-change Theorem 2.43
θ	the rotation angle between functions/images f_1, f_2, see e.g. the Rotation Theorem 2.45
$f_1 * f_2$	the convolution of functions f_1 and f_2, see Definition 2.58 or discrete periodic convolution of functions f_1 and f_2, see Definition 3.22
$\widetilde{f}(x,y), \widetilde{f_1}(x,y), \widetilde{f_2}(x,y)$	the periodization of functions $f(x,y)$, $f_1(x,y)$, $f_2(x,y)$: $\{0,1,\ldots,N-1\} \to \mathbb{C}, N \in \mathbb{N}$, see Definition 3.6
w	the number of possible pixel values of a gray-scale image, see Definition 4.1
Round	the unary operation of rounding to the nearest integer
$g_{GR}, g_{GC}, g_{HR}, g_{HC}$	Gaussian and Hanning, rectangular and circular window functions, see Definition 4.11
$f^p(\rho,\varphi)$	function $f(x,y)$ transformed to polar coordinates, see Definition 4.12
$f^{lp}(\rho,\varphi)$	function $f(x,y)$ transformed to logarithmic-polar coordinates, see Definition 4.14
$r_1, r_2, \sigma_1, \sigma_2$	parameters of the low-pass high-pass weight function, see Definition 4.19
λ	parameter of the Gaussian low-pass weight function, see Definition 4.20
ε	radius of the neighborhood used for subpixel precision registration, see Section 4.4
T_ς	tangential high-pass filter used for solar corona images, see Section 4.6.1

Appendix A

Proof of equality $\int_0^\infty \frac{\sin x}{x}\,\mathrm{d}x = \frac{\pi}{2}$

A.1 Integration-theory proof

Several times, we needed to compute the integral

$$\int_0^\infty \frac{\sin px}{x}\,\mathrm{d}x.$$

It is enough to compute the integral

$$\int_0^\infty \frac{\sin x}{x}\,\mathrm{d}x \tag{A.1}$$

as

$$\int_0^\infty \frac{\sin px}{x}\,\mathrm{d}x = \left|\begin{array}{rcl} t &=& px \\ \mathrm{d}t &=& p\,\mathrm{d}x \end{array}\right| = \int_0^\infty \frac{\sin t}{\frac{t}{p}}\frac{\mathrm{d}t}{p} = \int_0^\infty \frac{\sin t}{t}\,\mathrm{d}t.$$

I learned the computation of (A.1) at the lectures in Integration Theory by Christer Borell at Chalmers University of Technology in autumn 2009 [11]. It is also Exercise 59 on page 77 in [27], where Folland suggests to integrate $\mathrm{e}^{-xy}\sin x$ with respect to x and y.

Let us start with computing

$$\int_0^b \left(\int_0^\infty \mathrm{e}^{-xy}\sin x\,\mathrm{d}y\right)\,\mathrm{d}x, \qquad b \in \mathbb{R}^+.$$

If the iterated integral with the absolute value of integrand is finite, we can use Fubini's Theorem [9].

$$\int_0^b \left(\int_0^\infty |\mathrm{e}^{-xy}\sin x|\,\mathrm{d}y\right)\,\mathrm{d}x = \int_0^b \left[-\frac{|\sin x|}{x}\mathrm{e}^{-xy}\right]_{y=0}^\infty\,\mathrm{d}x = \int_0^b \frac{|\sin x|}{x}\,\mathrm{d}x \leq 1\cdot b < +\infty$$

Now we can use Fubini's Theorem in its weak form and write

$$\int_0^b \left(\int_0^\infty \mathrm{e}^{-xy} \sin x \, \mathrm{d}y \right) \mathrm{d}x = \int_0^\infty \left(\int_0^b \mathrm{e}^{-xy} \sin x \, \mathrm{d}x \right) \mathrm{d}y. \tag{A.2}$$

Now we compute

$$\begin{aligned}
\int \mathrm{e}^{-xy} \sin x \, \mathrm{d}x &= \int \mathrm{e}^{-xy} \left(\frac{1}{2\mathrm{i}} \mathrm{e}^{\mathrm{i}x} - \frac{1}{2\mathrm{i}} \mathrm{e}^{-\mathrm{i}x} \right) \mathrm{d}x = \frac{1}{2\mathrm{i}} \int \left(\mathrm{e}^{-xy+\mathrm{i}x} - \mathrm{e}^{-xy-\mathrm{i}x} \right) \mathrm{d}x = \\
&= \frac{1}{2\mathrm{i}} \left(\frac{\mathrm{e}^{-xy+\mathrm{i}x}}{-y+\mathrm{i}} - \frac{\mathrm{e}^{-xy-\mathrm{i}x}}{-y-\mathrm{i}} \right) = \frac{\mathrm{e}^{-xy}}{2\mathrm{i}} \left(\frac{\mathrm{e}^{\mathrm{i}x}}{-y+\mathrm{i}} - \frac{\mathrm{e}^{-\mathrm{i}x}}{-y-\mathrm{i}} \right) = \\
&= \frac{\mathrm{e}^{-xy}}{2\mathrm{i}} \frac{\mathrm{e}^{\mathrm{i}x}(-y-\mathrm{i}) - \mathrm{e}^{-\mathrm{i}x}(-y+\mathrm{i})}{y^2+1} = \\
&= \frac{\mathrm{e}^{-xy}}{2\mathrm{i}(y^2+1)} \left[-\mathrm{i} \left(\mathrm{e}^{\mathrm{i}x} + \mathrm{e}^{-\mathrm{i}x} \right) + y \left(-\mathrm{e}^{\mathrm{i}x} + \mathrm{e}^{-\mathrm{i}x} \right) \right] = \\
&= \frac{\mathrm{e}^{-xy}}{y^2+1} (-\cos x - y \sin x).
\end{aligned}$$

Therefore,

$$\begin{aligned}
\int_0^\infty \left(\int_0^b \mathrm{e}^{-xy} \sin x \, \mathrm{d}x \right) \mathrm{d}y &= \int_0^\infty \left[\frac{\mathrm{e}^{-xy}}{y^2+1} (-\cos x - y \sin x) \right]_{x=0}^b \mathrm{d}y = \\
&= \int_0^\infty \left(\frac{\mathrm{e}^{-by}}{y^2+1} (-\cos b - y \sin b) + \frac{1}{y^2+1} \right) \mathrm{d}y = \int_0^\infty f_b(y) \, \mathrm{d}y + \frac{\pi}{2}
\end{aligned}$$

if we denote

$$\frac{\mathrm{e}^{-by}}{y^2+1} (-\cos b - y \sin b) = f_b(y).$$

Furthermore,

$$\lim_{b \to \infty} f_b(y) = 0.$$

We need a majorant to $|f_b(y)|$ to use the Dominated Convergence Theorem [9].

$$|f_b(y)| \leq \frac{\mathrm{e}^{-by}}{y^2+1} (1+y) \leq \mathrm{e}^{-by} \left(\frac{1}{2} + 1 \right) = \frac{3}{2} \mathrm{e}^{-by}$$

as it is elementary to prove that for every $y \in \mathbb{R}$

$$\frac{2y}{y^2+1} \leq 1 \quad \text{and} \quad \frac{1}{y^2+1} \leq 1.$$

Therefore $f_b(y) \in \mathcal{L}^1$ with respect to Lebesque measure on $\langle 0, +\infty \rangle$ [9]. Using the Dominated Convergence Theorem, we obtain that

$$\lim_{b \to \infty} \int_0^\infty f_b(y) \, \mathrm{d}y = \int_0^\infty \lim_{b \to \infty} f_b(y) \, \mathrm{d}y = \int_0^\infty 0 \, \mathrm{d}y = 0.$$

The left-hand side of equation (A.2) is

$$\int_0^b \left[-\mathrm{e}^{-xy}\frac{\sin x}{x}\right]_{y=0}^{+\infty} \mathrm{d}x = \int_0^b \frac{\sin x}{x}\,\mathrm{d}x \to \int_0^\infty \frac{\sin x}{x}\,\mathrm{d}x \quad \text{as } b \to \infty.$$

The left-hand side of equation (A.2) is convergent to $\frac{\pi}{2}$, therefore the right-hand side is convergent as well and it holds

$$\int_0^\infty \frac{\sin x}{x}\,\mathrm{d}x = \frac{\pi}{2}. \tag{A.3}$$

A.2 Fourier-series proof

There is another, more basic proof of the equality (A.3), which I learned after having written down the nice proof above. In fact, it was an exam task at the exam from the course in Fourier Analysis at Chalmers University of Technology on October 24, 2009 prepared by Hjalmar Rosengren. The proof consists of two steps. First, we will show that

$$\frac{\sin x}{x} = b_0 + \sum_{n=1}^{\infty} b_n \cos(nx), \qquad 0 < x < \pi,$$

where

$$b_n = \frac{1}{\pi}\int_{(n-1)\pi}^{(n+1)\pi} \frac{\sin x}{x}\,\mathrm{d}x.$$

Then, using this result, we will compute the integral (A.1).

$\frac{\sin x}{x}$ is an even function, therefore it can be expressed as a Fourier cosine series [26]

$$\frac{\sin x}{x} = \frac{b_0}{2} + \sum_{n=1}^{\infty} b_n \cos(nx), \qquad 0 < x < \pi,$$

where

$$b_n = \frac{2}{\pi}\int_0^\pi \frac{\sin x}{x}\cos(nx)\,\mathrm{d}x. \tag{A.4}$$

A well-known trigonometric formula [38] says that for any real numbers A, B it holds

$$2\cos\frac{A+B}{2}\sin\frac{A-B}{2} = \sin A - \sin B.$$

If we take $\frac{A-B}{2} = x$, $\frac{A+B}{2} = nx$ and separate A, B, we find out that

$$2\sin x\cos(nx) = \sin(x(n+1)) - \sin(x(n-1)).$$

Plugging this into the formula for b_n (A.4), we get

$$b_n = \frac{1}{\pi}\int_0^\pi \left(\frac{\sin(x(n+1))}{x} - \frac{\sin(x(n-1))}{x}\right)\mathrm{d}x.$$

$$\frac{1}{\pi}\int_0^{\pi}\frac{\sin(x(n+1))}{x}\,\mathrm{d}x = \left|\begin{array}{ll} t = x(n+1) & 0 \mapsto 0 \\ \mathrm{d}t = (n+1)\,\mathrm{d}x & \pi \mapsto \pi(n+1)\end{array}\right| = \frac{1}{\pi}\int_0^{\pi(n+1)}\frac{\sin t}{t}\,\mathrm{d}t \qquad (A.5)$$

A similar procedure gives that

$$\frac{1}{\pi}\int_0^{\pi}\frac{\sin(x(n-1))}{x}\,\mathrm{d}x = \frac{1}{\pi}\int_0^{\pi(n-1)}\frac{\sin t}{t}\,\mathrm{d}t \qquad (A.6)$$

Taking the difference between formulae (A.5) and (A.6) gives the desired formula for b_n in the form

$$b_n = \frac{1}{\pi}\int_{(n-1)\pi}^{(n+1)\pi}\frac{\sin x}{x}\,\mathrm{d}x.$$

Plugging $n = 0$ into the last formula will make us use $\frac{\sin x}{x}$ for $x = 0$. The function value is undefined, one should rather use the notation of ζ functions introduced by Formulae (2.3) on page 10. However, this would make the following integrals too complicated. Therefore, from now on, we will assume that $\frac{\sin x}{x}$ equals 1 for $x = 0$. Finally, we can write $\frac{\sin x}{x}$ on $0 < x < \pi$ as

$$\frac{\sin x}{x} = \frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{\sin t}{t}\,\mathrm{d}t + \frac{1}{\pi}\sum_{n=1}^{\infty}\cos(nx)\int_{(n-1)\pi}^{(n+1)\pi}\frac{\sin t}{t}\,\mathrm{d}t.$$

If we now take $x = 0$, we calculate

$$\begin{aligned} 1 &= \frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{\sin t}{t}\,\mathrm{d}t + \frac{1}{\pi}\sum_{n=1}^{\infty}\int_{(n-1)\pi}^{(n+1)\pi}\frac{\sin t}{t}\,\mathrm{d}t = \\ &= \frac{1}{\pi}\int_0^{\pi}\frac{\sin t}{t}\,\mathrm{d}t + \frac{1}{\pi}\sum_{\substack{n=-\infty \\ n\neq 0}}^{\infty}\int_{n\pi}^{(n+1)\pi}\frac{\sin t}{t}\,\mathrm{d}t = \frac{1}{\pi}\int_{-\infty}^{\infty}\frac{\sin t}{t}\,\mathrm{d}t. \end{aligned}$$

This means that

$$\int_{-\infty}^{\infty}\frac{\sin t}{t}\,\mathrm{d}t = \pi$$

concluding that

$$\int_0^{\infty}\frac{\sin x}{x}\,\mathrm{d}x = \frac{\pi}{2}.$$

Appendix B

Proof of equality $\int_0^\infty \frac{|\sin x|}{x}\,\mathrm{d}x = \infty$

The proof is taken from [11].

$$\int_0^\infty \frac{|\sin x|}{x}\,\mathrm{d}x = \sum_{k=0}^{\infty} \int_{k\pi}^{(k+1)\pi} \frac{|\sin x|}{x}\,\mathrm{d}x \geq \sum_{k=0}^{\infty} \int_{k\pi+\frac{\pi}{4}}^{(k+1)\pi-\frac{\pi}{4}} \frac{|\sin x|}{x}\,\mathrm{d}x \geq$$

$$\sum_{k=0}^{\infty} \int_{k\pi+\frac{\pi}{4}}^{(k+1)\pi-\frac{\pi}{4}} \frac{|\sin x|}{(k+1)\pi-\frac{\pi}{4}}\,\mathrm{d}x = \sum_{k=0}^{\infty} \frac{1}{(k+1)\pi-\frac{\pi}{4}} \underbrace{\int_{\frac{\pi}{4}}^{\frac{3\pi}{4}} |\sin x|\,\mathrm{d}x}_{>0} = +\infty$$

Appendix C

Proof of equality $\int_0^\infty \frac{\sin^2 x}{x}\,\mathrm{d}x = \infty$

$$\int_0^\infty \frac{\sin^2 x}{x}\,\mathrm{d}x = \int_0^1 \frac{\sin^2 x}{x}\,\mathrm{d}x + \int_1^\infty \frac{\sin^2 x}{x}\,\mathrm{d}x$$

The first integral is a finite number. To show this we define a function

$$\psi(x) = \begin{cases} \frac{\sin^2 x}{x} & \text{if } x > 0 \\ 0 & \text{if } x = 0, \end{cases}$$

which is a non-negative function defined for non-negative numbers. The function is continuous in zero since $\lim_{x\to 0} \frac{\sin^2 x}{x} = 0$ by L'Hospital's rule. Then we can use Weierstrass Theorem [38] saying that a function which is continuous on a closed interval ($[0, 1]$ in this case) is also bounded on this interval and the integral from 0 to 1 is lower or equal to this upper bound. In fact, the function is bounded by 1 and the first integral is therefore smaller than one. Now we focus on the second integral.

$$\int_1^\infty \frac{\sin^2 x}{x}\,\mathrm{d}x = \int_1^\infty \frac{1-\cos 2x}{2x}\,\mathrm{d}x = \int_1^\infty \frac{1}{2x}\,\mathrm{d}x - \int_1^\infty \frac{\cos 2x}{2x}\,\mathrm{d}x \tag{C.1}$$

The first integral is a well-known diverging integral and it remains for prove that the second one converges. This can be proved using the Dirichlet criterion [17].

Theorem C.1 (Dirichlet criterion of integral convergence). If the following conditions are fulfilled,

1. function $\varphi(x)$ is monotone and converges to zero for $x \to +\infty$
2. function $f(x)$ has a bounded primitive function

$$F(x) = \int_a^x f(t)\,\mathrm{d}t,$$

then the integral

$$\int_a^\infty f(x)\varphi(x)\,\mathrm{d}x$$

converges (not necessarily absolutely).

Applying the Dirichlet criterion on the second integral in (C.1) taking $\varphi(x) = \frac{1}{2x}$, $f(x) = \cos 2x$ shows that the integral converges. The original integral

$$\int_0^\infty \frac{\sin^2 x}{x}\,\mathrm{d}x$$

was split in three integrals, two of which are finite and one diverges, therefore

$$\int_0^\infty \frac{\sin^2 x}{x}\,\mathrm{d}x = \infty.$$

Appendix D

Definitions of quadruple improper integral and $\mathcal{L}\left(\mathbb{R}^4\right)$

Definition D.1. Let $f(x,y,u,v)$ be a function $\mathbb{R}^4 \to \mathbb{R}$. We write

$$\lim_{(x,y,u,v)\to(\infty,\infty,\infty,\infty)} f(x,y,u,v) = L$$

if and only if for every $\varepsilon > 0$ there exists $r \in \mathbb{R}$ such that for every $a,b,c,d > r$: $|f(a,b,c,d) - L| < \varepsilon$.

Definition D.2 (Improper quadruple integral). Let $f(x,y,u,v)$ be a function $\mathbb{R}^4 \to \mathbb{R}$. Let $R = \langle a_1,\infty) \times \langle b_1,\infty) \times \langle c_1,\infty) \times \langle d_1,\infty)$, $a_1,b_1,c_1,d_1 \in \mathbb{R}$. If the following limits exist and are equal for every permutation of intervals and order of integration

$$\lim_{(a_1,b_1,c_1,d_1)\to(\infty,\infty,\infty,\infty)} \int_{a_1}^{a_2}\left(\int_{b_1}^{b_2}\left(\int_{c_1}^{c_2}\left(\int_{d_1}^{d_2} f(x,y,u,v)\,\mathrm{d}v\right)\mathrm{d}u\right)\mathrm{d}y\right)\mathrm{d}x = \cdots =$$

$$= \lim_{(a_1,b_1,c_1,d_1)\to(\infty,\infty,\infty,\infty)} \int_{d_1}^{d_2}\left(\int_{c_1}^{c_2}\left(\int_{b_1}^{b_2}\left(\int_{a_1}^{a_2} f(x,y,u,v)\,\mathrm{d}x\right)\mathrm{d}y\right)\mathrm{d}u\right)\mathrm{d}v = A$$

then we define

$$\iiiint_R f(x,y,u,v)\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}u\,\mathrm{d}v = A.$$

Analogically, the integral is defined for $R = (-\infty,a_2\rangle \times \langle b_1,\infty) \times \langle c_1,\infty) \times \langle d_1,\infty)$, $R = (-\infty,a_2\rangle \times (-\infty,b_2\rangle \times \langle c_1,\infty) \times \langle d_1,\infty)$ etc. Furthermore, if all the following integrals exist and are finite, or, in case some of the integrals are infinite, they have the

same sign

$$\begin{aligned}
\iiiint\limits_{\langle 0,\infty\rangle^4} f(x,y,u,v)\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}u\,\mathrm{d}v &= A_1\\
\iiiint\limits_{(-\infty,0\rangle\times\langle 0,\infty\rangle^3} f(x,y,u,v)\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}u\,\mathrm{d}v &= A_2\\
\iiiint\limits_{\langle 0,\infty\rangle\times(-\infty,0\rangle\times\langle 0,\infty\rangle^2} f(x,y,u,v)\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}u\,\mathrm{d}v &= A_3\\
&\vdots\\
\iiiint\limits_{(-\infty,0\rangle^4} f(x,y,u,v)\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}u\,\mathrm{d}v &= A_{16},
\end{aligned}$$

we define

$$\iiiint\limits_{(-\infty,\infty)^4} f(x,y,u,v)\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}u\,\mathrm{d}v = \sum_{n=1}^{16} A_n.$$

If function $f(x,y)$ is $\mathbb{R}^2 \to \mathbb{C}$, Definition 2.1 on page 5 remains valid and numbers $A_1, A_2, \ldots, A_{16}$ are complex.

Definition D.3 ($\mathcal{L}\,(\mathbb{R}^4)$ space). Let us denote $\mathcal{L}\,(\mathbb{R}^4)$ the space of all complex functions of four real variables, i.e.,

$$f(x,y,u,v) = p(x,y,u,v) + \mathrm{i}q(x,y,u,v),$$

where $p, q : \mathbb{R}^4 \to \mathbb{C}$ such that

$$\iiiint\limits_{\mathbb{R}^4} |f(x,y,u,v)|\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}u\,\mathrm{d}v$$

exists and is finite.

Convention D.4. Let $a_1, a_2, b_1, b_2, c_1, c_2, d_1, d_2 \in \mathbb{R} \cup \{-\infty, \infty\}$, $a_1 < a_2, b_1 < b_2, c_1 < c_2, d_1 < d_2$ and let $f(x,y,u,v) \in \mathcal{L}\,(\mathbb{R}^4)$. Due to Fubini's theorem [9] the quadruple integral of function f on the cartesian product $\langle a_1, a_2\rangle \times \langle b_1, b_2\rangle \times \langle c_1, c_2\rangle \times \langle d_1, d_2\rangle$ can equivalently be written in many forms, such as

$$\begin{aligned}
&\iiiint\limits_{\langle a_1,a_2\rangle\times\langle b_1,b_2\rangle\times\langle c_1,c_2\rangle\times\langle d_1,d_2\rangle} f(x,y,u,v)\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}u\,\mathrm{d}v =\\
&= \int_{a_1}^{a_2}\int_{b_1}^{b_2}\left(\int_{c_1}^{c_2}\int_{d_1}^{d_2} f(x,y,u,v)\,\mathrm{d}u\,\mathrm{d}v\right)\mathrm{d}x\,\mathrm{d}y =\\
&= \int_{a_1}^{a_2}\int_{b_1}^{b_2}\int_{c_1}^{c_2}\int_{d_1}^{d_2} f(x,y,u,v)\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}u\,\mathrm{d}v.
\end{aligned}$$