INSTITUT FÜR INFORMATIK

UND PRAKTISCHE MATHEMATIK

Low-Level Image Processing with the Structure Multivector

Michael Felsberg

Bericht Nr. 0203
15. März 2002

CHRISTIAN-ALBRECHTS-UNIVERSITÄT

KIEL

Institut für Informatik und Praktische Mathematik der
Christian-Albrechts-Universität zu Kiel
Olshausenstr. 40
D – 24098 Kiel

Low-Level Image Processing with the Structure Multivector

Michael Felsberg

Bericht Nr. 0203
15. März 2002

e-mail: mfe@isy.liu.se

“Dieser Bericht gibt den Inhalt der Dissertation wieder, die der Verfasser
im Dezember 2001 bei der Technischen Fakultät der
Christian-Albrechts-Universität zu Kiel eingereicht hat.
Datum der Disputation: 12. Februar 2002”

1. Gutachter	Prof. G. Sommer (Kiel)
2. Gutachter	Prof. U. Heute (Kiel)
3. Gutachter	Prof. J. J. Koenderink (Utrecht)

Datum der mündlichen Prüfung: 12.2.2002

ABSTRACT

The present thesis deals with two-dimensional signal processing for computer vision. The main topic is the development of a sophisticated generalization of the one-dimensional analytic signal to two dimensions. Motivated by the fundamental property of the latter, the invariance – equivariance constraint, and by its relation to complex analysis and potential theory, a two-dimensional approach is derived. This method is called the monogenic signal and it is based on the Riesz transform instead of the Hilbert transform. By means of this linear approach it is possible to estimate the local orientation and the local phase of signals which are projections of one-dimensional functions to two dimensions. For general two-dimensional signals, however, the monogenic signal has to be further extended, yielding the structure multivector. The latter approach combines the ideas of the structure tensor and the quaternionic analytic signal. A rich feature set can be extracted from the structure multivector, which contains measures for local amplitudes, the local anisotropy, the local orientation, and two local phases. Both, the monogenic signal and the structure multivector are combined with an appropriate scale-space approach, resulting in generalized quadrature filters. Same as the monogenic signal, the applied scale-space approach is derived from the three-dimensional Laplace equation instead of the diffusion equation. Hence, the two-dimensional generalization of the analytic signal turns out to provide a whole new framework for low-level vision. Several applications are presented to show the efficiency and power of the theoretic considerations. Among these are methods for orientation estimation, edge and corner detection, stereo correspondence and disparity estimation, and adaptive smoothing.

ACKNOWLEDGMENT

First of all, I appreciate my supervisors Prof. Gerald Sommer and Prof. Ulrich Heute for guiding me through the interesting fields of computer vision and signal theory. They allowed me to work as independently as was necessary to obtain substantial new results. However, without the countless, intensive discussions with Prof. Gerald Sommer and without his scientific intuition I would not have been able to develop the presented ideas and to write this thesis as it is. He supported me whenever it was necessary.

Furthermore, I acknowledge the support and help of all our former and present members of the cognitive systems group in Kiel. Especially the fruitful discussions with Norbert Krüger, Thomas Bülow, and Hongbo Li have had a great influence on my work. I thank Gerd Diesner and Henrik Schmidt for their technical support and Françoise Maillard for her help in administrative matters.

I am grateful for all the discussions with other scientists and for the references they have given to me, especially Kieran Larkin, Robin Owens, Terence Tao, Dieter Betten, Joachim Weickert, Christoph Schnörr, Bernd Jähne, Hanno Scharr, Wolfgang Förstner, Gösta Granlund, Josef Bigün, Moshe Porat, Nir Sochen, Tony Lindeberg, and Jan Koenderink.

I acknowledge Hongbo Li, Jim Byrnes, Norbert Krüger, and Christian Perwass for proofreading parts of this thesis.

I am also grateful to the German National Merit Foundation (Studienstiftung des deutschen Volkes) and the DFG Graduiertenkolleg No. 357 for their financial support during my PhD studies.

Last, but not least, I thank my family for mental support, my parents Hans and Beate, my sister Susanne, my parents-in-law Otfried and Rosemarie, my brother-in-law Sebastian, my daughter Lisa-Marie and especially my wife Regina, to whom I dedicate this thesis.

CONTENTS

Chapter 1

INTRODUCTION

> *'This question is best answered by making*
> *use of one quite singular improper rotation,*
> *namely reflection in O;*
> *it carries any point P into its antipode P' ...'*
>
> Hermann Weyl (1885-1955)

The quotation above[1] which states that the answer is a reflection in the origin can be considered as *the* central point of my thesis. However, what is the question to this answer?

This questions could be: *'Is there a sophisticated 2D generalization of the Hilbert transform which can be used for image processing and if there is, what does it look like?'*

As it can be concluded from the question above, this thesis is about a topic which is related to three different fields, as there are mathematics (harmonic analysis), engineering (signal theory), and computer science (image processing). The intention of this thesis is to close a gap between the three different areas in order to make a theory, which is yet widely unknown in signal theory and image processing, accessible for applications.

1.1 Terms and Motivation

Image processing is a quite wide field in general. In this thesis, I focus on *low-level image processing* which is the part of image processing being closest to the signal layer. Hence, it is the first layer in a *bottom-up* system design [60]. Low-level image processing is partially based on 2D signal theory which means that *grayscale images* (I do not consider color images in this thesis) are interpreted as

[1] This quotation is borrowed from a talk of K. Larkin. All other quotations are taken from [145].

sampled and quantized 2D intensity functions on finite domains. Although images are discrete signals, the signal theory which is dealt with in this thesis is formulated for the continuous case. From my point of view, it is more reasonable to consider the *underlying continuous signal* obtained from the discrete signal by interpolation rather than the discrete signal itself. This is a fundamental difference between 1D and nD signal processing which can be motivated by considering the sampling process. Whereas 1D signals can only be discretized by sampling the line (non-uniformly in general), 2D signals can be discretized according to *infinitely many different grids*, since any (non-uniform) grid which covers the plane can be used for sampling. Basically no grid is preferable and the optimal sampling depends on the context – if *shifts* of the signal are to be considered, a uniform, rectangular grid is optimal, whereas for *dilations* and *rotations* a (log-)polar grid is optimal (see e.g. [122, 130, 141]). Covering all possible mentioned *group actions* with one grid at a time is *impossible*. This can only be achieved in the case of continuous signals.

Considering continuous signal processing, it is desirable to have *linear* and *invariant* methods. Linearity is advantageous since *stability*, i.e., bounded input bounded output (BIBO), and *robustness*, i.e., flexibility according to small deviations from the underlying *model*, are often impossible to prove for non-linear systems. Especially the use of *thresholds* turns nearly every stable and robust system into an unstable and non-robust system which follows immediately from the discontinuous character of a threshold. However, thresholds are needed for introducing *symbols* which are elementary to classical *knowledge based approaches*. Thus, the requirement for stable and robust approaches imply not to use knowledge based approaches, but *appearance based approaches*. On the other hand, purely appearance based approaches are closely related to unbiased bottom-up strategies of system design which is not reasonable unless certain constraints are introduced [135].

Hence, one of the aims of this thesis is to present linear signal preprocessing methods which are designed under a specific point of view, i.e., their output represents a *continuous measure* of a certain *feature*. This feature should not be an abstract or even meaningless entity, but a *semantically interpretable property* of the signal (see also [80]). This requirement for the design of the linear preprocessing layer can be considered as a necessary *bias* which allows to develop a capable medium-level layer or at least simplifies its development. Although it is sometimes misleading to draw parallels to the human visual system, the biased design of stable preprocessing layers is consistent with experiments from neuro-physiology which have shown that certain regions in the visual cortex can be modeled as Gabor filters [77] and hence, these regions can be considered to measure feature probabilities (see also [87]).

On the other hand, features are worthless unless they are *invariant* under irrelevant changes of the signal. Invariance means that a feature value is not changed by a certain group acting on a signal. The opposite to invariance is *equivariance*, which

means that there is a monotonic dependency of a feature value on the parameter of the group action. If a set of features includes only invariant and equivariant features, it fulfills the *invariance – equivariance* property. This property is especially useful in signal analysis since it allows to estimate signal properties independently of each other [61]. If a set of features fulfills the invariance – equivariance requirement and is at the same time a unique description of the signal, it performs a *split of identity*. In the split of identity different features represent mutually different properties of the signal and the whole set of features is a complete description of the signal.

The probably most well known *global* description of a signal which meets the split of identity requirement is the Fourier transform. The Fourier transform of a signal at an arbitrary position can be decomposed into the Fourier transform of the signal located at the origin and a complex *modulation factor* which determines the shift of the signal. The complete decomposition into a unique signal representation and a signal independent representation of shifts can also be generalized to other group operations (see e.g. [141, 51]). However, group actions on *local parts* of signals are not covered by these global approaches, and therefore it is reasonable to have localized spectral properties, i.e., *local phase* and *local amplitude*[2]. The local phase and the local amplitude are the spatial domain counterparts of the phase spectrum and the amplitude spectrum, respectively. It is well known that the phase spectrum of a signal covers most of the *structural information* of the signal [111]. The structure of a signal can be defined as the local decomposition of the signal according to its symmetries. From a simplified point of view, the signal is decomposed into impulses (even symmetry) and jumps (odd symmetry) which is reflected by the local phase. The amplitude spectrum however, represents different signal properties, e.g. the energy of a signal. Hence, the local amplitude can be used to estimate the local energy of a signal. Local structure and local energy are independent information which perform a split of identity (at least for signals with a sufficient narrow bandwidth, see [61], page 171, and below).

In 1D signal processing the appropriate tool for estimating the local amplitude and the local phase is the *analytic signal*. Basically, the analytic signal is the linear combination of a signal and its *Hilbert transform*. The method of the analytic signal is a quite universal tool which can be applied in many different fields as there are geophysics [103], radar signal processing [107], airfoil design [147], etc.. However, these applications are all out of the scope of this thesis and I will focus on structure analysis of images. The fundamental property of the analytic signal for this purpose is its symmetry which is induced by the symmetry of the Hilbert transform. The idea of the Hilbert transform is to create a signal which is related to the orig-

[2] According to [91], the modulus of a frequency response is called *magnitude* response, whereas the modulus of a spectrum is called *amplitude* spectrum. Since 'magnitude spectrum' and 'amplitude response' also occur in the literature [133], using one or the other term seems to be rather a matter of taste. However, I prefer the term amplitude to be consistent the term local *amplitude*.

inal signal by being in *quadrature*. Considering the Fourier basis functions solely, the Hilbert transform of a cosine function yields the corresponding sine function and vice versa. Hence, the Hilbert transform is basically a phase shift by $\pi/2$. In general, signals contain different spectral components with different amplitudes and phases. Although the Hilbert transform of such a 'frequency cocktail' is still its $\pi/2$ phase shifted version, the straight forward interpretation and the independence of local amplitude and local phase is lost. Due to the superposition of all frequency components, the information contained in each component is reduced to some mean value which is dominated by the high-amplitude frequency terms.

In order to 'demix' the 'frequency cocktail', the different frequency components of the signal must be separated. This can be done by introducing bandpass filters which select only a small part of the spectral information of the signal, and therefore allow to identify the 'ingredients' of the 'frequency cocktail'. The combination of a bandpass filter with its Hilbert transform yields a *quadrature filter* which consists of an even part (the bandpass filter) and an odd part (the Hilbert transform of the bandpass filter). Quadrature filters can be used to estimate the local amplitude and the local phase for a specific frequency range, and therefore they improve the structure analysis by distinguishing different frequency components.

The frequency of a signal is related to the *scale* of a signal since frequencies are reciprocal to spatial distances and distances change with scale. On a coarser scale the distance between two positions is smaller than on a finer scale. Considering a coarser scale means to filter a signal with a lowpass filter which eliminates high frequency components. Thus, bandpass filters are directly related to *scale-space* since they represent the difference of the signal representations at two distinct scales. From this point of view, quadrature filters fulfill the invariance – equivariance property with respect to local amplitude, local phase, and *local scale*. Although a change in phase and a change in scale seem to have totally different effects to a set of quadrature filter responses, the difference in the representation of both local properties is solely of algebraic nature. Since the phase is represented as a unit complex number, a shift is represented by a complex rotation, i.e., a complex modulation. Opposed to that, the scale of a signal is reflected by the responses of a set of bandpass filters, which possibly constitute a *frame*, i.e., a redundant 'basis'. Accordingly, a *dilation* (change in scale) moves the coefficients of the filters to other frequency components which is equivalent to a transformation on the space spanned by the bandpass filters, see also figure 1.1.

The different algebraic representations of phase-shifts and dilations are adequate since the phase is a periodic feature whereas the scale is not. Although this distinction seems to be obvious, many approaches in image processing do not take care of the adequate algebraic representation of features. According to [81], a famous, but symptomatic example of *inadequate* feature representation is the embedding of grayscale images as 2D surfaces in 3D Euclidean space for photometric

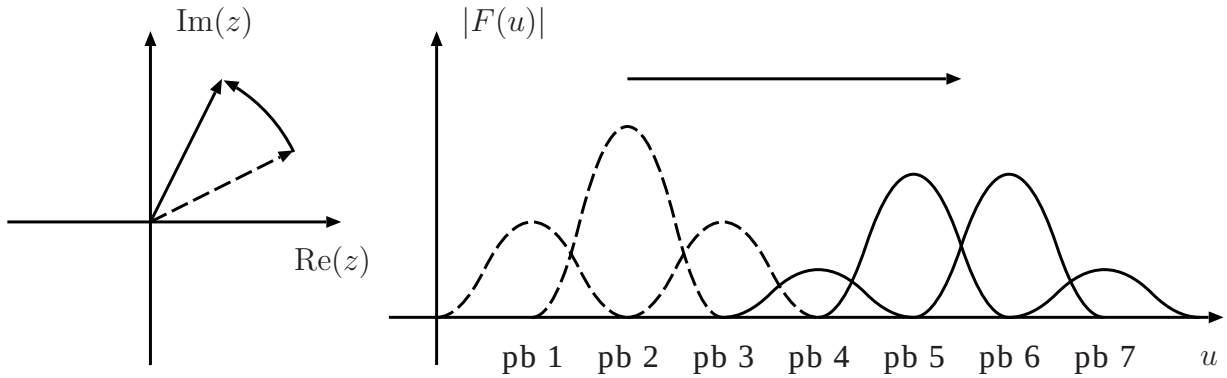

Fig. 1.1: Left: a shift of the signal results in a complex modulation of the filter response. Right: a dilation of the signal moves the coefficients to other passbands (pb).

applications. Rotations in that 3D space rather lead to absurd results than to relevant photometric operations. Whereas the grayscales are given by a single-valued function, the frequency components of a signal are given by a many-valued function, and therefore the representation as a *set* of bandpass coefficients is adequate. Such a representation is also known as the *channel representation* of a feature [60, 52], which can thus be considered as a specific algebraic embedding.

The restriction to certain, inadequate algebraic embeddings complicates the design of sophisticated approaches, especially in multi-dimensional signal processing. By using an adequate representation, many problems can be solved much easier and some non-linear problems become linear [135]. A quite famous example for a so far unsolved problem is the analytic signal and quadrature filters for multiple dimensions. For example, in [82] it is stated that no odd filters with isotropic energy exist and hence, no quadrature filters which are invariant under rotations can be developed. Since the odd part of a quadrature filter is obtained from the Hilbert transform of the even part of a quadrature filter (the bandpass), the previous statement is basically identical to the conjecture in [10] which claims that the Hilbert transform cannot be extended to higher dimensions in an isotropic way. The latter statement is based on a proof that the *dispersion relation* in 2D has no solutions [110]. The 1D dispersion relation can be used to derive the Hilbert transform.

The failure of developing an isotropic quadrature filter is caused by the restriction to complex numbers. If this restriction is omitted, a straightforward solution *can* be found, which is a central topic of this thesis. Considering the problem of a 2D analytic signal more in detail, it becomes obvious why the complex numbers are *insufficient* to represent the local phase in 2D. The complex analytic signal allows to represent two degrees of freedom: the local amplitude and the local 1D symmetry. If a 2D signal is generated from a 1D signal by identifying the 1D coordi-

nate with the scalar product of the 2D coordinate and an arbitrary unit vector, i.e., $f_2(x, y) = f_1(x \cos\theta + y \sin\theta)$, a new degree of freedom is introduced: the *direction* θ. Hence, the 2D signal has three degrees of freedom whereas a complex number can only represent two of them. The complex representation is therefore not complete and cannot fulfill the invariance – equivariance constraint.

The deeper analysis of the previous argumentation leads to even more fundamental concepts. The class of signals which are defined according to the construction above are obviously a small subset of general 2D signals. The characteristic property of this subset is that its elements are *intrinsically 1D*, which means that they only vary with respect to one *orientation*. The orientation of an intrinsically 1D signal is defined by the 1D subspace which is orthogonal to the constant lines in the signal. The direction of an intrinsically 1D signal however, cannot be defined by means of the signal solely. Further knowledge about the signal, e.g. about its formation (i.e., the involved unit vector), is necessary. Thus, the difference between orientation and direction is the *directional sense* of the latter, i.e., the orientation is an angle in $[0, \pi)$ whereas the direction is an angle in $[0, 2\pi)$, and therefore the orientation is given by the direction modulo π. Since the direction of a signal cannot be obtained from the signal locally (see also [61]), the best possible generalized 2D analytic signal yields an *oriented phase*. Oriented phase means that for intrinsically 1D (i1D) signals the phase is a feature with two degrees of freedom describing the 1D symmetry and the orientation of the symmetry axis. Such a phase approach represents the *i1D 2D phase*. Most of the classical approaches to a 2D analytic signal (see next section) deal with this kind of phase approach.

The class of truly 2D signals, i.e., 2D signals which do not vary with respect to one orientation only, cannot be described by a phase with one degree of freedom at all. The space of 2D symmetries is *infinite* times larger than the space of 1D symmetries which directly follows from considering the Radon transform of a 2D signal. The *Radon transform* projects a 2D signal onto an oriented line so that an orientation parameterized space of 1D functions is obtained. Each of these 1D functions may have its own 1D symmetry and since the orientation parameter is continuous (at least in theory), a 2D symmetry is decomposed into infinitely many 1D symmetries. Hence, the complete set of 2D symmetries cannot be represented by a finite number of features. In practice, it is reasonable to assume that the angular variation of a signal is smooth to some degree which can also be interpreted as a *band limitation* of the angular spectrum. Since the angle is a periodic coordinate, the angular spectrum is discrete and a band limitation implies a finite number of Fourier coefficients. However, this number of coefficients is much too large for applying algorithms with reasonable time complexity. One has to reduce the number of coefficients further so that a *compromise* between computational effort and sufficient discrimination of slightly different symmetries is obtained.

For the design of filters in general, the angular variation of signals has led to

different approaches. One of the most sophisticated solutions is the principle of *steerable filters* [57, 132], which is based on the shift theorem of the Fourier series. In combination with the assumption that the number of relevant Fourier coefficients is limited (see above), a flexible and computationally effective approach is obtained. Steerable filters are based on a number of basis filters which are applied to the signal and whose responses are superposed in order to obtain the final filter output. The method of basis filters can be generalized [146, 102] and it can be combined with an orientation adaptive scheme [2, 64], yielding a *semi-linear*, shift-variant preprocessing. Semi-linear means that the linearity constraint is fulfilled if two signals with the same orientation are added, but it is violated if their orientations are different. Most approaches to steerable filters are designed to obtain a scalar valued filter output.

Another way to cope with the angular variation of signals is the *structure tensor* [7, 55]. The basic idea of the structure tensor is to restrict the number of angular coefficients to second order terms and to create a tensor valued filter response. In spite of the restriction to second order terms, the discriminative power of the structure tensor has been proved to be sufficient in practical applications. The structure tensor provides an adequate method for representing the amplitudes of i2D signals, but cannot be considered as a generalized analytic signal since it is non-linear and does not contain a phase approach.

Besides the new 2D generalization of the analytic signal for i1D signals, the design of a sophisticated 2D analytic signal for i2D signals is a main topic of this thesis. Such a generalized analytic signal is obtained by combining the ideas of steerable filters, the structure tensor, and the quaternionic analytic signal (see next section). The algebraic embedding of the new approach as the *structure multivector* yields five independent local signal features: local (main) orientation, local amplitude, local isotropy, and two local phases in the classical sense. If the considered signal is i1D, the structure multivector is automatically reduced to the i1D approach.

Quite surprisingly, the derivation of the i1D approach also leads to a new linear *scale-space* which differs from the Gaussian scale-space. The new scale-space allows to combine the new generalized analytic signals with bandpass filters which emerge from the same underlying theory. By introducing local metrics induced by the signal, the scale-space becomes adaptive which leads to optimal estimates of the image features. Hence, the search for a generalized analytic signal has resulted in a new generic framework for low-level image processing.

Parts of this work have already been published in [46, 44, 48, 43, 49, 50, 45, 47]. Those ideas, derivations, definitions, theorems, and algorithms in chapter 3–6 which are not indicated as taken from literature and which are no standard mathematical methods are original contributions of the author.

1.2 Known Approaches for the 2D Analytic Signal

In the literature, several approaches for generalizing the analytic signal to two dimensions can be found. Most of these approaches make use of the complex numbers. The only exception I know is the *quaternionic analytic signal* [17, 19] which tries to encode local 2D symmetries using three phase angles.

Most approaches to a generalized analytic signal are based on 2D generalizations of the Hilbert transform. The simplest extension of the Hilbert transform to 2D is to apply the 1D Hilbert transform with respect to one of the coordinate axes. Hence, the 2D analytic signal is obtained by omitting either the left half-space or the lower half-space of the Fourier domain. To see the drawbacks of this approach, it is necessary to recall the definition of the 1D analytic signal. Although the DC-component of a signal is irrelevant for the analysis of its structure, it distorts the phase in a non-linear manner. Consider, for example, the oscillation $\cos(2\pi x)$. Adding a DC of value one does not affect the Hilbert transform of the signal. Hence, the analytic signal reads

$$1 + \exp(i2\pi x) = 2\exp(i\pi x)\cos(\pi x) \ .$$

As it can be observed from this example, the phase can be changed totally just by introducing a DC-component. Actually, this is no problem in the 1D case since the DC-component of the signal can easily be suppressed. This is not true for the 2D extension described above. The DC-component is now replaced by a xDC-component (or yDC-component) which means that those parts of the signal which do not differ in x (or y) direction act as the DC-component in the 1D case, i.e., they simply add to the analytic signal, and therefore they distort the phase in a non-linear manner. Accordingly, these *partial Hilbert transforms* are totally insufficient, if there is any energy on the frequency axis which is perpendicular to the direction of the Hilbert transform.

One can try to overcome this problem by rotating the partial Hilbert transform so that the critical axis lies in an area without energy. This area defines a *preference direction* for the Hilbert transform and herewith for the analytic signal. This preference direction has to be chosen by a further preprocessing step and it depends on the local neighborhood of the signal. In general, there is no global optimal orientation. Furthermore, instead of one filter set, a number of filter sets (depending on the angular sampling) has to be designed and as many convolutions have to be performed as there are different optimal orientations. Hence, the computational load is significantly increased. Nevertheless, phase distortions might occur, since the Hilbert transforms of some parts of the signal might nearly cancel out each other, because they lie on different sides of the 'DC-line' resulting in different signs of their local phases. Consider, for example, the signal $f(x, y) = \cos(x + y) + \cos(x - y)$ (the 'egg-tray signal', see figure 1.2) and the correspond-

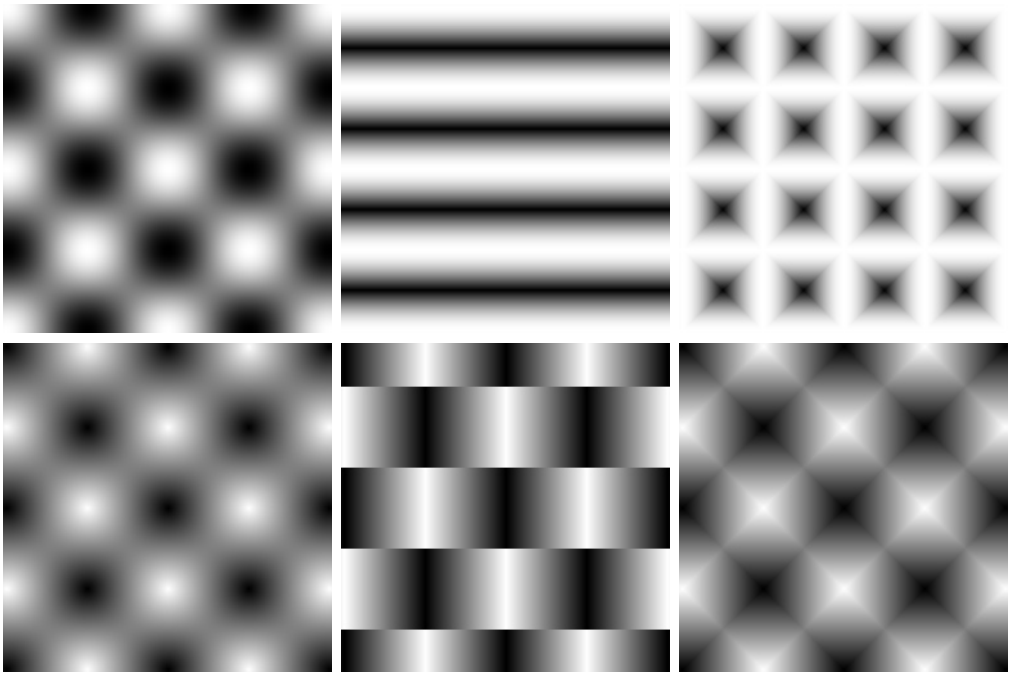

Fig. 1.2: Upper row (from left to right): original egg-tray signal (white corresponds to the value two and black corresponds to the value minus two), local amplitude obtained from the horizontal Hilbert transform, and local amplitude obtained from the partial Hilbert transform with preference direction. In the two amplitude images white corresponds to the value two and black corresponds to the value zero. Bottom row (from left to right): original phase, local phase obtained from the horizontal Hilbert transform, and local phase obtained from the partial Hilbert transform with preference direction. In the phase images white corresponds to the value π and black corresponds to the value zero.

ing correct local phase[3] (see also figure 1.2). The horizontal partial Hilbert transform yields $f_H(x,y) = -\sin(x+y) - \sin(x-y)$. Hence, the analytic signal reads $f(x,y) - if_H(x,y) = 2\exp(ix)\cos(y)$. The corresponding (incorrect) phase and amplitude can be found in the center column of figure 1.2. Obviously, the y-part of the oscillation is not extended to a complex signal. This is not a DC-caused problem, but due to the half-space symmetry of the partial Hilbert transform. Amplitude and phase distortions also occur in the case of choosing adaptively the preference direction, where the latter is obtained from the gradient of the signal. For the present example, the gradient is given as

$$-(\sin(x+y)+\sin(x-y),\, \sin(x+y)-\sin(x-y))^T = -(\sin(x)\cos(y),\, \cos(x)\sin(y))^T \ .$$

[3] For the following discussion it is sufficient to consider the absolute value of the phase.

Since the considered signal consists of four Dirac impulses in the frequency domain, the partial Hilbert transform needs only be computed for four different cases. Taking into account solely the absolute value of the phase, the partial Hilbert transform of the signal reduces to

$$f_H(x,y) = -2 \begin{cases} \sin(x+y) + \sin(x-y) & \text{if } \sin(x)\cos(y) \geq \cos(x)\sin(y) \\ \sin(x+y) - \sin(x-y) & \text{if } \sin(x)\cos(y) < \cos(x)\sin(y) \end{cases} .$$

The local phase and the local amplitude obtained from this orientation adaptive Hilbert transform can be found in the right column of figure 1.2.

These amplitude and phase distortions are always present unless the signal only varies with respect to one direction, i.e., it is intrinsically one dimensional (see section 3.4 for a formal definition of the intrinsic dimension). For the latter case, the partial Hilbert transform results in the correct solution if the preference direction is chosen properly with respect to the orientation of the signal.

Due to its quite high computational load, the partial Hilbert transform with preference direction is commonly replaced by one of the two following methods:

1. Sampling of the orientation. The analytic signal is computed for a fixed set of angles. Due to the fact that the energy on the critical line is not doubled, the analytic signal with the highest energy is probably the one with the least phase distortions (which is not certain). If the orientation is sampled with a frequency[4] of $\pi/2$, exactly the same analytic signal is obtained for the egg-tray signal as in the case of the ideal partial Hilbert transform with preference direction. Hence the phase and amplitude distortions are the same (see figure 1.2).

2. The Hilbert transform is approximated by *steerable filters* [57, 2] (see section 3.5). This decreases the computational load, but introduces further phase distortions compared to the partial Hilbert transform with preference direction since the Hilbert transform is only approximated by a Fourier series (see also section 3.5).

All approaches to a partial Hilbert transform with preference direction have to be orientation adaptive in some way, and therefore they are semi-linear and computationally more expensive than a purely linear approach. Depending on the method, amplitude and phase distortions vary, but are always present for intrinsically 2D signals. In general, a lower computational load of the method goes in hand with larger distortions of phase and amplitude.

[4] Sampling the orientation with $\pi/2$ yields the Hilbert transform with respect to x and y, i.e., two real valued convolutions must be performed which is the same computational load as for the new approach (see chapter 4).

Besides the partial Hilbert transform, there exists the *total Hilbert transform* which is defined by the subsequent application of the partial Hilbert transforms with respect to the two coordinate axes. The resulting frequency response is real-valued and symmetric with respect to the origin. Hence, the total Hilbert transform does not perform a phase shift of $\pi/2$, i.e., $\cos(\cdot)$ and $\sin(\cdot)$ are no longer a Hilbert pair [17]. The analytic signal formed by this approach does not have a local phase which can be related to the 1D phase, and therefore it is not a valid 2D generalization of the analytic signal with respect to the split of identity.

Motivated by the one-sidedness of the 1D analytic signal in frequency domain, Hahn suggested a *single-orthant analytic signal* [65, 66] which means that the 2D analytic signal is obtained by omitting all spectral components which have at least one negative coordinate, i.e., only the first quadrant is kept. The resulting approach is not complete and for taking into account all signal parts it is necessary to have a second analytic signal. The latter is obtained by keeping either the second or the fourth quadrant. Considering both analytic signals at a time, a complete signal representation is obtained. According to [17], this representation is not satisfactory unless an appropriate algebraic embedding is chosen.

The most sophisticated approach to a 2D analytic signal which occurred in the literature so far is the *quaternionic analytic signal*. It combines the ideas of partial and total Hilbert transforms and the one-orthant spectrum. Due to the algebraic embedding, the quaternionic analytic signal is complete. The quaternionic Fourier transform is four-fold redundant which means that for a real signal any of the quadrants in the quaternionic Fourier domain contains the whole information of the signal[5]. The main drawback of the quaternionic analytic signal is the missing isotropy resulting from the fact that it is fixed to the coordinate axes. For the egg-tray signal, the quaternionic analytic signal yields a correct solution: $f_{\mathrm{Q}}(x, y) = 2\exp(ix)\exp(jy)$. Amplitude and phase are obtained according to [17] and are basically correct. The amplitude has the value two throughout the spatial domain and the phases (the quaternionic analytic signal yields three phase angles) are obtained as $(\phi, \psi, \theta) = (x_{[-\pi,\pi)}, 0, y_{[-\pi/2,\pi/2)})$ which contains the separated original phase information[6] (see figure 1.3). The non-separated local phase (see figure 1.2, bottom left) can be recovered by the function $\arccos(\cos\phi\cos\theta)$. The drawback of the quaternionic analytic signal is however, that it only works perfectly if the signal and the coordinate axes are compatible with respect to their orientations. If the egg-tray signal is rotated, the quaternionic signal also fails to recover a constant local amplitude (see figure 1.3) and a correct local phase.

Finally, all previously known 2D generalizations of the analytic signal fail to estimate the correct amplitude and phase of intrinsically 2D signals. Although the

[5] The redundancy is obviously lost if the signal is not real valued as in [124], where Sangwine uses the quaternions to encode color information.

[6] The subscript of $x_{[-\pi,\pi)}$ means that x is obtained modulo the interval $[-\pi, \pi)$, accordingly for y.

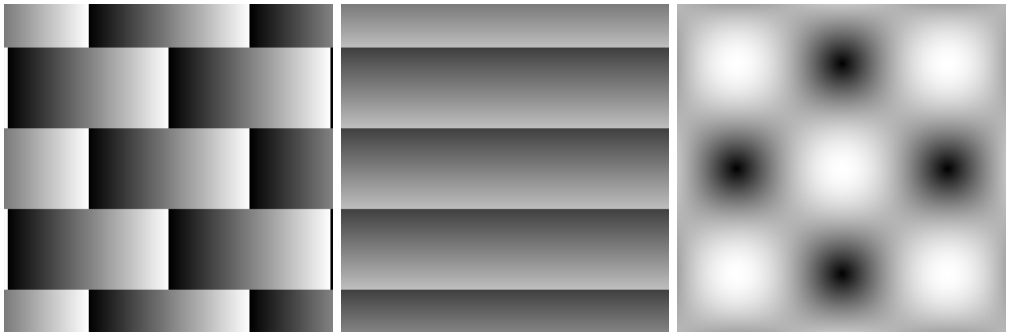

Fig. 1.3: Left and center: ϕ- and θ-phase of the quaternionic analytic signal for the egg-tray signal (white corresponds to π and black corresponds to $-\pi$). The ψ-phase and the local amplitude are constant for compatible orientations. Right: local amplitude for incompatible orientation (egg-tray signal rotated by $\pi/4$, black means amplitude zero, white means amplitude two).

quaternionic analytic signal is quite close to a correct solution, it suffers from its anisotropic behavior. As it turns out, the rotation invariance is the key to an optimal solution.

1.3 Outline of this Thesis

In the second chapter, a brief introduction to geometric algebra of 2D and 3D Euclidean space is given. The introduction covers also group representations and Clifford analysis on the introduced algebras. The material which is presented is self-contained and it is reduced to the relevant part which will be needed in the remaining chapters of this thesis.

Using the introduced algebraic framework, parts of classical signal theory are reformulated in the third chapter. Besides LSI operators and the Fourier transform for 1D and 2D signals, the 1D analytic signal and its phase are defined in detail. Further sections deal with the idea of the structure tensor, the intrinsic dimension, and steerable filters.

The fourth chapter is the central chapter of this thesis. Motivated by a vector field based derivation of the Hilbert transform, the Riesz transform is derived from the 3D Laplace equation. The Riesz transform is the 2D generalized Hilbert transform yielding the monogenic signal as a generalized analytic signal. Based on the Radon transform an i1D 2D phase approach is derived and the orientation – phase decomposition is discussed. Besides the generalized analytic signal, a new linear scale-space is established by the Laplace equation. The new scale-space yields a

natural design principle for quadrature filters, which can be interpreted as a continuous wavelet transform.

In chapter five, two extensions of the monogenic signal are discussed. By introducing a local metric, an anisotropic scale-space is obtained which can be computed efficiently by a first order Taylor expansion and steerable filters. Having anisotropic, steerable quadrature filters concludes the work on i1D signals. For i2D signals an extension of the monogenic signal is formulated which yields a seven component multivector for the description of 2D symmetries. This extended approach allows to estimate the amplitude, the main orientation, the isotropy, and two phases of i2D neighborhoods, i.e., five independent features.

The sixth chapter is about applications of the presented theoretic framework. Besides optimal orientation estimation and contrast independent edge detection, methods for stereo correspondence, disparity estimation, adaptive smoothing, and corner detection are presented.

Each of the chapters two to six are concluded by a short summary which allows the reader to recall the main topics. Chapter seven concludes this thesis. It contains a summary of the presented results, a discussion on open problems, and possible topics of future work. The thesis is completed by an appendix, the list of references, and a subject index.

Chapter 2

GEOMETRIC ALGEBRA

> *'I regard it as an inelegance, or imperfection, in quaternions, or rather in the state to which it has been hitherto unfolded, whenever it becomes or seems to become necessary to have recourse to x, y, z, etc..'*
>
> Sir William Rowan Hamilton (1805-1865)

In this thesis all derivations and signal processing approaches are formulated in *geometric algebra* for the following reasons:

1. In the introduction I have shown that the algebra of complex numbers is *insufficient* to formulate and derive a sophisticated 2D (or nD) generalization of the analytic signal.

2. I make use of vectors and group operations in this thesis, where the former are commonly denoted by column vectors and the latter are commonly represented by matrix algebra. Furthermore, signal theory is closely related to complex analysis which introduces the complex numbers as a representation of 2D vectors, 2D rotations, and scale transformations. Hence, different mathematical approaches for representing the same entities are mixed in the classical formulation of multidimensional signal theory.[1]

3. Group representations by means of matrix algebra are not intuitive with respect to their geometric interpretation. The rotation angle and the rotation axis of a 3D rotation do not obviously show up in their corresponding matrix, nor is it straightforward to recognize its inverse rotation matrix.

4. Classical vector algebra does not distinguish between *axial vectors* and *polar vectors* [63], i.e., angles and positions in Euclidean space are represented in the same way.

[1] This can be observed in e.g. [8].

The mixture of the two kinds of vectors, complex numbers, and matrices makes it very hard to formulate coherent derivations in the classical framework. One has to *switch between different representations* in order to be able to perform necessary calculations or to interpret the results. For example, a rotation can be represented by a matrix or by an axis and an angle, forming a *rotation vector*. While the matrix is needed to calculate the effect of the rotation to a vector, the rotation vector is the most compact representation of a rotation [115] and at the same time it gives an interpretation of the group action. Since the matrix representing a rotation does not directly show the axis and the angle (except for trivial cases), it is hard to switch between the two representations.[2] Another example is the identification of complex numbers with 2D vectors on one hand, and with *scaling-rotations* (rotations and isotropic scale transformations) in the plane on the other hand. In order to perform certain calculations, both identifications are mixed, and it does not become evident from the formula if a complex number in a calculation represents a 2D vector or a scaling-rotation operation. Due to these *inadequacies* of the commonly known mathematical methods, I decided to use *geometric algebra* of Euclidean space throughout this thesis. I hope that I will convince the reader of the adequacy of this mathematical framework while keeping the thesis not too technical in the mathematical sense.

In the first two sections of this chapter, I will give simple introductions to the geometric algebra of 2D and 3D Euclidean space, which are limited to those aspects being relevant for understanding this thesis. I have often met people raising objections to geometric algebra, who hold the opinion that it is unfavorable to use such 'complicated mathematics' in an engineering topic. Actually, I think that working with geometric algebra is *much easier* than using the classical framework and I hope to convince the reader of this fact in the remaining sections and chapters. Already in the third and fourth section of this chapter, the elegance of representations in geometric algebra becomes evident. In section 2.3 the connection between complex analysis and harmonic 2D vector fields is directly obtained from the embedding into geometric algebra. As a consequence, the higher dimensional generalization of complex analysis is obtained in a straightforward way by means of harmonic 3D vector fields. Hence, the highly complicated and abstract theory of Clifford analysis is easily obtained from geometric concepts.

For a more general introduction to geometric algebra (based on a physical background), see [70], for a quite complete overview, see [71]. Much easier to understand are the tutorials [36, 35] which are both available online. The GABLE tutorial [36] is combined with a Matlab[3]-package for numerical computations using geometric algebra. Some other more mathematical pieces of work are [1, 97] and [34, 118] (with a focus on group theory). For a more detailed introduction to Clifford analy-

[2] For 3D this mapping from matrices to angle and axis is quite complicated, see [24], page 318.
[3] Matlab is a registered trademark of The MathWorks, Inc.

sis I refer to [13, 98]. At this point, I also want to refer to the difference between *geometric* algebra and *Clifford* algebra. Whereas the latter is the more traditional, mathematical term, the former indicates basically the same algebra, but with focus on geometric applications (see introduction of [6]).

2.1 Geometric Algebra of $\mathbb{R}^2$

In the beginning of this chapter, the identification of complex numbers with 2D vectors on the one hand and with 2D scaling-rotations on the other hand has already been mentioned. These distinct interpretations are now used to introduce the geometric algebra of 2D Euclidean space $\mathbb{R}^2$.

2.1.1 Definition of the Algebra

The Euclidean space $\mathbb{R}^2$ is spanned by the orthonormal basis vectors $\mathbf{e}_1$ and $\mathbf{e}_2$, see figure 2.1. Orthonormality means that $\mathbf{e}_i \cdot \mathbf{e}_j = \delta_{ij}$ where the dot indicates the scalar

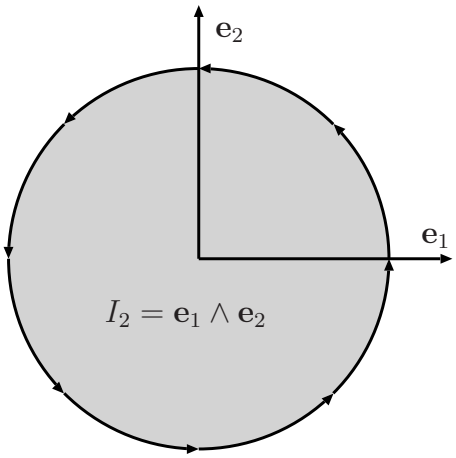

Fig. 2.1: 2D Euclidean space: basis vectors $\mathbf{e}_1$, $\mathbf{e}_2$, and the unit pseudoscalar I_2 which represents the rotation from $\mathbf{e}_1$ to $\mathbf{e}_2$.

product and δ_{ij} indicates the Kronecker delta. For every two vectors $\mathbf{x} = x\mathbf{e}_1 + y\mathbf{e}_2$ and $\mathbf{u} = u\mathbf{e}_1 + v\mathbf{e}_2$ the *dot product* is identical to the scalar product[4] and is given by distributivity and associativity:

$$\mathbf{x} \cdot \mathbf{u} = xu\,\mathbf{e}_1 \cdot \mathbf{e}_1 + xv\,\mathbf{e}_1 \cdot \mathbf{e}_2 + yu\,\mathbf{e}_2 \cdot \mathbf{e}_1 + yv\,\mathbf{e}_2 \cdot \mathbf{e}_2 = xu + yv \ . \tag{2.1}$$

[4] The dot product is later generalized to other algebra elements.

Note that vectors are always indicated in bold face, whereas their components are in italics. Furthermore, I circumvent indices in the components, because otherwise triple indices would occur later. In geometric algebra vectors are not represented by column or row vectors, as it is common in matrix algebra, but solely as linear combinations of basis elements. All calculations are performed as for polynomials (i.e., applying associativity and distributivity) with the difference that the products between the basis elements follow different rules (e.g. some are not commutative, see below).

The 2D Euclidean space is a real vector space, i.e., the field of the vector space is given by the real numbers. Hence, there is a third basis element in the algebra, the scalar 1. The multiplication of vectors by a scalar $\lambda \in \mathbb{R}$ is well known since it is just the *scalar multiplication* $\lambda\mathbf{x} = \lambda x\mathbf{e}_1 + \lambda y\mathbf{e}_2$. Obviously, the scalar multiplication differs from the dot product, and hence two different products exist. Going back to the interpretation of a complex number as a scaling-rotation, it is natural to identify the scalar multiplication with the scale transformation. But how to represent the rotation? In order to obtain a rotation, a basis element is needed which changes $\mathbf{e}_1$ to $\mathbf{e}_2$ (i.e., a rotation by 90°). The same basis element should change $\mathbf{e}_2$ to $-\mathbf{e}_1$. Indicating this basis element by I_2 (see figure 2.1), the previous observation yields that $I_2^2 = I_2 I_2 = -1$ (rotation by 180°), and hence I_2 cannot be obtained by linear combination of the first three basis elements. In order to represent I_2 in terms of its effect on the basis vectors, a new product is introduced, called the *wedge product*:

$$I_2 \stackrel{\text{def}}{=} \mathbf{e}_1 \wedge \mathbf{e}_2 = -\mathbf{e}_2 \wedge \mathbf{e}_1 \stackrel{\text{def}}{=} \mathbf{e}_{12} \ . \tag{2.2}$$

The wedge product is anti-commutative, which follows from the interpretation of I_2 as a rotation by 90°. Due to the interpretation as a rotation by 90°, the imaginary unit i of the complex numbers can be substituted by I_2, yielding an *algebra isomorphism*, i.e., every calculation in the complex numbers can be performed by a calculation in the algebra of the space spanned by $\{1, I_2\}$ with the product defined by the relations above. Since both spaces are linear 2D spaces, the only thing which is left to show to be equal is the multiplication:

$$(a + bI_2)(c + dI_2) = ac + adI_2 + bcI_2 + bdI_2I_2 = (ac - bd) + (ad + bc)I_2 \ .$$

Applying the algebra isomorphism, a rotation by angle θ can be represented by $\cos\theta + \sin\theta I_2 = \exp(\theta I_2)$ according to $\cos\theta + \sin\theta i = \exp(\theta i)$ in the complex numbers.

Now, having four basis elements $\{1, \mathbf{e}_1, \mathbf{e}_2, I_2 = \mathbf{e}_{12}\}$ it is possible to represent both vectors and scaling-rotations, but in contrast to the complex numbers, the representations of both are strictly disjunct. Therefore, a rotation is not applied to a vector by means of 'complex' multiplications or the dot product. The previously used multiplications have to be extended to obtain the *algebra product* or *geometric product* which is simply denoted by concatenation of the factors. This extension is obtained by interpreting a vector as the rotated and scaled basis vector $\mathbf{e}_1$:

$\lambda(\cos\theta\mathbf{e}_1 + \sin\theta\mathbf{e}_2) = \lambda\mathbf{e}_1\exp(\theta\mathbf{e}_{12})$. Hence, $\mathbf{e}_1\mathbf{e}_{12} = \mathbf{e}_2$ and $\mathbf{e}_2\mathbf{e}_{12} = -\mathbf{e}_1$, where it is important that $\mathbf{e}_{12}$ always multiplies from the right. In order to know the other products, i.e., $\mathbf{e}_{12}\mathbf{e}_1$ and $\mathbf{e}_{12}\mathbf{e}_2$, one needs to know the geometric product of two basis vectors. This product represents a relation between the two vectors (i.e., a transformation) and therefore does not contain vector components. Obviously, the square of a basis vector $\mathbf{e}_j\mathbf{e}_j$ is symmetric, therefore a scalar (the wedge is anti-symmetric) and identical to the dot product. The geometric product $\mathbf{e}_1\mathbf{e}_2$ has no scalar components since the dot product is zero, and accordingly, it is identical to the wedge product. Summarizing the geometric products of the basis elements, the multiplication table 2.1 is obtained. Applying associativity and distributivity, the geometric

	1	$\mathbf{e}_1$	$\mathbf{e}_2$	$\mathbf{e}_{12}$
1	1	$\mathbf{e}_1$	$\mathbf{e}_2$	$\mathbf{e}_{12}$
$\mathbf{e}_1$	$\mathbf{e}_1$	1	$\mathbf{e}_{12}$	$\mathbf{e}_2$
$\mathbf{e}_2$	$\mathbf{e}_2$	$-\mathbf{e}_{12}$	1	$-\mathbf{e}_1$
$\mathbf{e}_{12}$	$\mathbf{e}_{12}$	$-\mathbf{e}_2$	$\mathbf{e}_1$	-1

Tab. 2.1: Multiplication table of the basis elements $\{1, \mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_{12}\}$. The left factors are indicated by the first row and the right factors by the first column.

products of the basis elements establish also the geometric product of two arbitrary elements of the algebra:

$$\begin{aligned}(a_1 &+ b_1\mathbf{e}_1 + c_1\mathbf{e}_2 + d_1\mathbf{e}_{12})(a_2 + b_2\mathbf{e}_1 + c_2\mathbf{e}_2 + d_2\mathbf{e}_{12}) \\ &= (a_1a_2 + b_1b_2 + c_1c_2 - d_1d_2) + (a_1b_2 + b_1a_2 - c_1d_2 + d_1c_2)\mathbf{e}_1 \\ &+ (a_1c_2 + b_1d_2 + c_1a_2 - d_1b_2)\mathbf{e}_2 + (a_1d_2 + b_1c_2 - c_1b_2 + d_1a_2)\mathbf{e}_{12}\ .\end{aligned} \tag{2.3}$$

The associative algebra formed by the four basis elements $\{1, \mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_{12}\}$ and their geometric product is called the *geometric algebra of 2D Euclidean space* and is denoted by $\mathbb{R}_2$.[5]

Having defined the geometric algebra $\mathbb{R}_2$, some further notations are introduced now. The element obtained by a linear combination of wedge products of basis vectors (in 2D there is only one such product: $\mathbf{e}_1 \wedge \mathbf{e}_2$) is called a *bivector*. Furthermore, geometric algebras are *graded algebras*, which means that scalars are elements of *grade* zero, vectors are of grade one, and bivectors are of grade two. A general combination of elements with different grade is called a *multivector*:

$$M = a + b\mathbf{e}_1 + c\mathbf{e}_2 + d\mathbf{e}_{12}\ . \tag{2.4}$$

The k-grade part of a multivector M is obtained from the *grade operator*: $\langle M\rangle_k$. Hence, $\langle M\rangle_0$ is the scalar part of M, $\langle M\rangle_1$ is the vector part of M (and as such an element of $\mathbb{R}^2$), and $\langle M\rangle_2$ is the bivector part of M.

[5] The most common notation is $\mathbb{R}_{2,0}$, but I skip the second index since I am only working on Euclidean spaces.

By means of the geometric product the dot product and the wedge product can be extended to arbitrary multivectors, introducing the *inner product* and the *outer product*, respectively. The geometric product of two vectors can be decomposed into their dot product and their wedge product:

$$\mathbf{xu} = \mathbf{x}\cdot\mathbf{u} + \mathbf{x}\wedge\mathbf{u} = xu + yv + (xv - yu)\mathbf{e}_{12} \ ,$$

and therefore dot and wedge product of vectors can also be defined according to the commutator and the anti-commutator of the geometric product

$$\mathbf{x}\cdot\mathbf{u} = \frac{1}{2}(\mathbf{xu} + \mathbf{ux}) \ \text{ and} \tag{2.5}$$

$$\mathbf{x}\wedge\mathbf{u} = \frac{1}{2}(\mathbf{xu} - \mathbf{ux}) \ . \tag{2.6}$$

The geometric product of a scalar and an arbitrary multivector is given by the scalar multiplication which is *by definition* identical with the outer product. The geometric product of a bivector with a vector or a bivector is reduced to their inner product which is especially interesting in the context of *duality*. The unit bivector of $\mathbb{R}_2$, $\mathbf{e}_{12} = I_2$, is also called the *unit pseudoscalar* indicating the positive directional sense.[6] The geometric product of a multivector M by I_2^{-1} from the right is called the *dual* of M, and it is indicated by a star:

$$M^* = MI_2^{-1} = -M\mathbf{e}_{12} \ .$$

The modulus of a multivector $M = M_0 + M_1\mathbf{e}_1 + M_2\mathbf{e}_2 + M_{12}\mathbf{e}_{12}$ is given by its Euclidean norm (the *magnitude* according to [71]):

$$|M| = \sqrt{M_0^2 + M_1^2 + M_2^2 + M_{12}^2} \ . \tag{2.7}$$

Hence, for any vector $\mathbf{x}$, its modulus is obtained from the square root of the product with itself, and therefore its inverse vector is given by $\mathbf{x}^{-1} = \frac{\mathbf{x}}{\mathbf{x}\cdot\mathbf{x}}$. The modulus of a multivector can also be obtained by means of the *reversion* and the grade operator. The reversion is denoted by a tilde and it is defined by

$$\tilde{M} = \langle M\rangle_0 + \langle M\rangle_1 - \langle M\rangle_2 \ , \tag{2.8}$$

and hence

$$|M| = \sqrt{\langle M\tilde{M}\rangle_0} \ . \tag{2.9}$$

The reversion describes the result when the basis vectors are multiplied in the reverse order, therefore nothing happens for scalars and vectors, but bivectors are negated. The reversion is an *involutive antiautomorphism*, which means that

$$\tilde{\tilde{M}} = M \quad \text{and} \quad (MN)^{\sim} = \tilde{N}\tilde{M} \ . \tag{2.10}$$

The reversion plays the role of the complex conjugate in many cases.

[6] Obviously, $-I_2$ is also a unit pseudoscalar representing a rotation from $\mathbf{e}_2$ to $\mathbf{e}_1$, i.e., a negative rotation. In the sequel, the term 'the unit pseudoscalar' denotes the unit pseudoscalar with positive directional sense. Since each nD space has such a unit pseudoscalar, the index denotes the dimension of the underlying vector space $\mathbb{R}^n$, in order to prevent ambiguities.

2.1.2 Group Representation in $\mathbb{R}_2$

After having introduced all necessary algebraic terms, I now focus on the group representation. The combination of a scalar and a bivector, if nonzero, is equivalent to a scaling-rotation, and it is called a *spinor* [63]. For every spinor $s = r\exp(\theta I_2)$, the inverse spinor is given by

$$s^{-1} = r^{-1}\exp(-\theta I_2) = \frac{\tilde{s}}{|s|^2} \quad . \tag{2.11}$$

All spinors form a proper subalgebra of $\mathbb{R}_2$, the even-grade subalgebra $\mathbb{R}_2^+$. As mentioned above, $\mathbb{R}_2^+$ and $\mathbb{C}$ are isomorphic as algebras. Complex numbers which were interpreted as scaling-rotations, are now represented as spinors: $z = x + iy = r\exp(i\theta)$ is replaced by $s_{\mathbf{x}} = x + yI_2 = r\exp(\theta I_2)$. Opposite to that, complex numbers which were interpreted as vectors, are represented as the action of spinors to the basis vector $\mathbf{e}_1$: $\mathbf{x} = \mathbf{e}_1 s_{\mathbf{x}} = x\mathbf{e}_1 + y\mathbf{e}_2$, where $s_{\mathbf{x}} = \mathbf{e}_1\mathbf{x}$.

If the scaling-rotation is interpreted in terms of *Lie groups* [125], the *generator* of the rotation is the unit pseudoscalar I_2. The group of scale transformations is generated by the scalar 1. Computing the exponential map of $\lambda + \theta I_2$ proves that 1 and I_2 are the generators of the mentioned groups: $\exp(\lambda + \theta I_2) = e^{\lambda}\exp(\theta I_2)$. The generators are obtained by the inverse of the exponential map, which is obtained from the complex logarithm and the isomorphism $\mathbb{R}_2^+ \simeq \mathbb{C}$. The logarithm plays an important role for the definition of the phase (see section 3.2). Therefore, an explicit definition is given at this place (see also [70], page 75): The inverse of the exponential map $\exp : \mathbb{R}_2^+ \to \mathbb{R}_2^+$ is given by $\log : \mathbb{R}_2^+ \to \mathbb{R}_2^+$ where

$$\log(s) = \log(|s|) + \frac{\langle s\rangle_2}{|\langle s\rangle_2|}\operatorname{atan}\left(\frac{|\langle s\rangle_2|}{\langle s\rangle_0}\right) \tag{2.12}$$

for $s \in \mathbb{R}_2^+$, $\log(|s|)$ being the ordinary real logarithm of the modulus of s, and atan being the arc tangent mapping for the interval $[0, \pi)$. The $\mathbb{R}_2^+$-logarithm defined by (2.12) is the isomorphic mapping to the complex logarithm. Since the generator of a rotation is the bivector I_2 and since the rotation by the angle θ is represented by θI_2, one can think of angles as bivectors. Although this interpretation is not necessary in 2D, it becomes natural in 3D.

Besides the scaling-rotations in 2D, *reflections* with respect to an arbitrary line through the origin can also be expressed in an algebraic way. Consider a unit vector $\mathbf{n}$ indicating the orientation of the line of reflection (see figure 2.2). The vector $\mathbf{x}$ which is reflected with respect to that line can be decomposed into a part that is parallel to the line and a part that is orthogonal to the line. The former part is indicated by $\mathbf{x}^{\parallel}$ and the latter part is indicated by $\mathbf{x}^{\perp}$. Hence, $\mathbf{x} = \mathbf{x}^{\parallel} + \mathbf{x}^{\perp}$ and the reflected point is given by $\mathbf{x}' = \mathbf{x}^{\parallel} - \mathbf{x}^{\perp}$. The part $\mathbf{x}^{\parallel}$ is obviously given by projecting

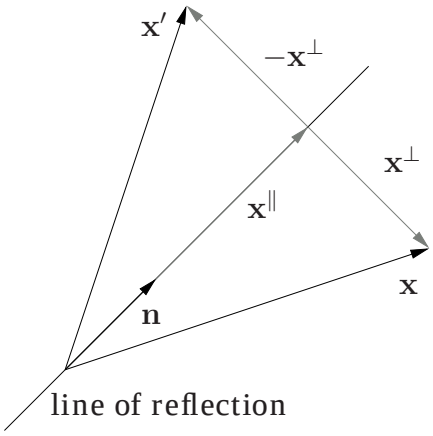

Fig. 2.2: Reflection in 2D. The vector $\mathbf{n}$ indicates the orientation of the line of reflection and $\mathbf{x}$ is the vector which is reflected onto $\mathbf{x}'$.

$\mathbf{x}$ onto $\mathbf{n}$, i.e., $\mathbf{x}^\parallel = (\mathbf{x}\cdot\mathbf{n})\mathbf{n}$, which is called the *projection* of $\mathbf{x}$ (with respect to $\mathbf{n}$). The remaining part of $\mathbf{x}$ is given by

$$\mathbf{x}^\perp = \mathbf{x} - \mathbf{x}^\parallel = \mathbf{x} - (\mathbf{x}\cdot\mathbf{n})\mathbf{n} = (\mathbf{x}\mathbf{n})\mathbf{n} - (\mathbf{x}\cdot\mathbf{n})\mathbf{n} = (\mathbf{x}\mathbf{n} - \mathbf{x}\cdot\mathbf{n})\mathbf{n} = (\mathbf{x}\wedge\mathbf{n})\mathbf{n} \qquad (2.13)$$

and it is called the *rejection* of $\mathbf{x}$ (with respect to $\mathbf{n}$). Finally, the reflected point $\mathbf{x}'$ is obtained as

$$\mathbf{x}' = \mathbf{x}^\parallel - \mathbf{x}^\perp = (\mathbf{x}\cdot\mathbf{n})\mathbf{n} - (\mathbf{x}\wedge\mathbf{n})\mathbf{n} = (\mathbf{n}\cdot\mathbf{x})\mathbf{n} + (\mathbf{n}\wedge\mathbf{x})\mathbf{n} = \mathbf{n}\mathbf{x}\mathbf{n}\ . \qquad (2.14)$$

Hence, any reflection with respect to a line through the origin is represented by the two-sided multiplication with the orientation vector.

At this point, I want to make two remarks. First, the reflection of the basis vector $\mathbf{e}_1$ yields the orientation vector with double argument, i.e., if $\mathbf{n} = \cos\theta\mathbf{e}_1 + \sin\theta\mathbf{e}_2$, $\mathbf{n}\mathbf{e}_1\mathbf{n} = \cos(2\theta)\mathbf{e}_1 + \sin(2\theta)\mathbf{e}_2$. This corresponds to the square of a unit complex number representing an orientation. The second remark is about the relation between reflections and rotations. Reflections are more fundamental since they can be used to define rotations. Two successive reflections, i.e., $\mathbf{n}_2\mathbf{n}_1\mathbf{x}\mathbf{n}_1\mathbf{n}_2$, correspond to the rotation about the origin by two times the angle between $\mathbf{n}_1$ and $\mathbf{n}_2$: $\mathbf{x}' = (\mathbf{n}_2\mathbf{n}_1)\mathbf{x}(\mathbf{n}_1\mathbf{n}_2) = \mathbf{x}(\mathbf{n}_1\mathbf{n}_2)^2 = \mathbf{x}(\cos\alpha + \sin\alpha\mathbf{e}_{12})^2 = \mathbf{x}\exp(2\alpha\mathbf{e}_{12})$ where α is the angle between $\mathbf{n}_1$ and $\mathbf{n}_2$.

Finally, the properties of the algebra $\mathbb{R}_2$ can be summarized as follows:

- $\mathbb{R}_2$ contains the vector space $\mathbb{R}^2$ including its scalar product which represents the collinear part of two vectors.
- The vector space is extended by the wedge product which represents the directed plane spanned by two vectors.

- $\mathbb{R}_2$ contains the complex numbers as a subalgebra. This subalgebra of spinors represents the group action of scaling-rotations on $\mathbb{R}^2$.
- The geometric product is the algebra product of the 4D algebra $\mathbb{R}_2$ by which all previously mentioned products can be expressed.
  - Applied to vectors, it represents the relation of two vectors.
  - Applied to spinors, it represents the group multiplication of scaling-rotations.
  - Applied to a vector and a spinor, it represents the result of the scaling-rotation group acting on the vector.
  - Applied as a two-sided multiplication with a unit vector, it represents the reflection with respect to the line given by the unit vector.

The advantages of geometric algebra are even more significant in the case of 3D vector space.

2.2 Geometric Algebra of $\mathbb{R}^3$

In this section the algebra of $\mathbb{R}^3$ and the group representation of 3D scaling-rotations are introduced.

2.2.1 Definition of the Algebra

The 3D Euclidean space $\mathbb{R}^3$ is spanned by the orthonormal basis $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ (see figure 2.3). Again, the scalar product of the basis vectors is given by the Kronecker delta: $\mathbf{e}_i \cdot \mathbf{e}_j = \delta_{ij}$, and therefore the dot product for arbitrary vectors $\mathbf{x} = x\mathbf{e}_1 + y\mathbf{e}_2 + z\mathbf{e}_3$ and $\mathbf{u} = u\mathbf{e}_1 + v\mathbf{e}_2 + w\mathbf{e}_3$ is defined by associativity and distributivity as $\mathbf{x} \cdot \mathbf{u} = xu + yv + zw$.

In the light of section 2.1, the vector space $\mathbb{R}^3$ is extended to form an algebra. At first, the wedge product of the basis vectors now yields three basis bivectors (see figure 2.3), and since the basis vectors are orthogonal, the wedge product of pairwise different basis vectors is equivalent to their geometric product :

$$\mathbf{e}_{12} \stackrel{\text{def}}{=} \mathbf{e}_1 \wedge \mathbf{e}_2 = \mathbf{e}_1\mathbf{e}_2 \tag{2.15}$$

$$\mathbf{e}_{23} \stackrel{\text{def}}{=} \mathbf{e}_2 \wedge \mathbf{e}_3 = \mathbf{e}_2\mathbf{e}_3 \tag{2.16}$$

$$\mathbf{e}_{31} \stackrel{\text{def}}{=} \mathbf{e}_3 \wedge \mathbf{e}_1 = \mathbf{e}_3\mathbf{e}_1 \ . \tag{2.17}$$

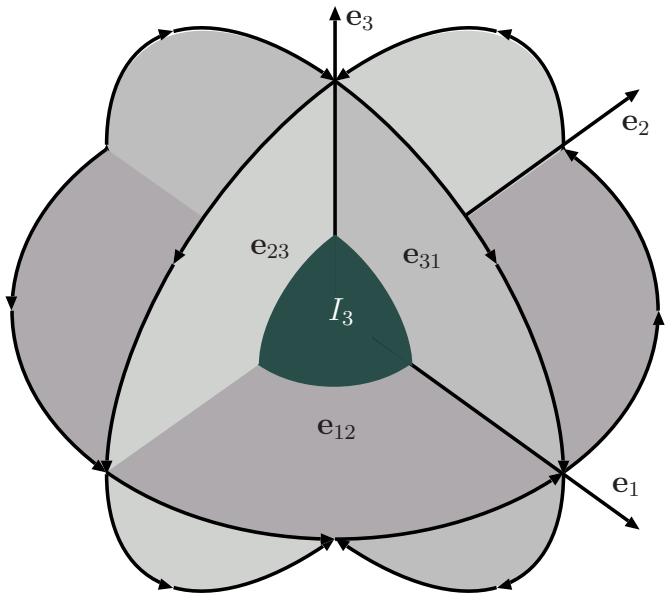

Fig. 2.3: 3D Euclidean space: the basis vectors $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$, the basis bivectors $\mathbf{e}_{12}$, $\mathbf{e}_{23}$, $\mathbf{e}_{31}$, and the unit pseudoscalar I_3.

In the case of $\mathbb{R}_2$ (see section 2.1) the geometric product of a bivector and a vector results in a vector (e.g. $\mathbf{e}_{12}\mathbf{e}_1 = -\mathbf{e}_2$). Now having three independent basis vectors, the geometric product of a basis bivector and the basis vector not contained in the bivector (e.g. $\mathbf{e}_{12}\mathbf{e}_3$) forms a new element, the basis *trivector* or unit pseudoscalar (see also figure 2.3):

$$I_3 \stackrel{\text{def}}{=} \mathbf{e}_{12}\mathbf{e}_3 = \mathbf{e}_1\mathbf{e}_2\mathbf{e}_3 \stackrel{\text{def}}{=} \mathbf{e}_{123} \quad . \tag{2.18}$$

From the definition of the pseudoscalar and the anti-commutativity of the product of pairwise different basis vectors, the other possible products forming the algebra are obtained. All products between the eight basis elements are summarized in table 2.2. From this multiplication table the geometric product of arbitrary multivectors is given by associativity and distributivity. Note that the pseudoscalar commutes with all other elements of the algebra.

Besides the geometric product, some other products which are used in this thesis are defined. These definitions again require the *grade operator* which selects that part of a multivector with a specific grade. If M is a general multivector in $\mathbb{R}_3$, then $\langle M \rangle_0$ is its scalar part, $\langle M \rangle_1$ is its vector part, $\langle M \rangle_2$ is its bivector part, and $\langle M \rangle_3$ is

	1	$\mathbf{e}_1$	$\mathbf{e}_2$	$\mathbf{e}_3$	$\mathbf{e}_{23}$	$\mathbf{e}_{31}$	$\mathbf{e}_{12}$	I_3
1	1	$\mathbf{e}_1$	$\mathbf{e}_2$	$\mathbf{e}_3$	$\mathbf{e}_{23}$	$\mathbf{e}_{31}$	$\mathbf{e}_{12}$	I_3
$\mathbf{e}_1$	$\mathbf{e}_1$	1	$\mathbf{e}_{12}$	$-\mathbf{e}_{31}$	I_3	$-\mathbf{e}_3$	$\mathbf{e}_2$	$\mathbf{e}_{23}$
$\mathbf{e}_2$	$\mathbf{e}_2$	$-\mathbf{e}_{12}$	1	$\mathbf{e}_{23}$	$\mathbf{e}_3$	I_3	$-\mathbf{e}_1$	$\mathbf{e}_{31}$
$\mathbf{e}_3$	$\mathbf{e}_3$	$\mathbf{e}_{31}$	$-\mathbf{e}_{23}$	1	$-\mathbf{e}_2$	$\mathbf{e}_1$	I_3	$\mathbf{e}_{12}$
$\mathbf{e}_{23}$	$\mathbf{e}_{23}$	I_3	$-\mathbf{e}_3$	$\mathbf{e}_2$	-1	$-\mathbf{e}_{12}$	$\mathbf{e}_{31}$	$-\mathbf{e}_1$
$\mathbf{e}_{31}$	$\mathbf{e}_{31}$	$\mathbf{e}_3$	I_3	$-\mathbf{e}_1$	$\mathbf{e}_{12}$	-1	$-\mathbf{e}_{23}$	$-\mathbf{e}_2$
$\mathbf{e}_{12}$	$\mathbf{e}_{12}$	$-\mathbf{e}_2$	$\mathbf{e}_1$	I_3	$-\mathbf{e}_{31}$	$\mathbf{e}_{23}$	-1	$-\mathbf{e}_3$
I_3	I_3	$\mathbf{e}_{23}$	$\mathbf{e}_{31}$	$\mathbf{e}_{12}$	$-\mathbf{e}_1$	$-\mathbf{e}_2$	$-\mathbf{e}_3$	-1

Tab. 2.2: Multiplication table of the basis elements $\{1, \mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_{23}, \mathbf{e}_{31}, \mathbf{e}_{12}, I_3\}$. The left factors are indicated by the first row and the right factors by the first column.

its pseudoscalar part. Using this notation, the following products are defined:

- The *inner product* is indicated by a dot. It is defined by

$$M \cdot N = \langle MN \rangle_{|r-s|} \tag{2.19}$$

  where M and N are two multivectors such that $M = \langle M \rangle_r$, $N = \langle N \rangle_s$, and $r, s > 0$. To be more specific, the following cases occur:

  - If either M or N is a scalar ($r = s = 0$), the inner product is zero.[7]
  - For two vectors ($r = s = 1$), the inner product is identical to their scalar product, defined as in the 2D case (2.5).
  - For a vector $\mathbf{x}$ ($r = 1$) and a bivector B ($s = 2$), the inner product is given by the vector part of the geometric product:

$$\mathbf{x} \cdot B = \langle \mathbf{x}B \rangle_1 = \frac{1}{2}(\mathbf{x}B - B\mathbf{x}) = -B \cdot \mathbf{x} \ . \tag{2.20}$$

  - The inner product of a vector $\mathbf{x}$ ($r = 1$) and a pseudoscalar λI_3 ($s = 3$) is given by their geometric product since $\langle \mathbf{x} I_3 \rangle_2 = \mathbf{x} I_3$:

$$\mathbf{x} \cdot (\lambda I_3) = \lambda \mathbf{x} I_3 = (\lambda I_3) \cdot \mathbf{x} \ . \tag{2.21}$$

  - The inner product of two bivectors B and C ($r = s = 2$) is given by

$$B \cdot C = \langle BC \rangle_0 = \frac{1}{2}(BC + CB) = C \cdot B \ . \tag{2.22}$$

[7] If one factor in MN is a scalar, the geometric product is just the scalar multiplication. The scalar multiplication of M by λ is represented by the outer product $\lambda \wedge M = M \wedge \lambda$, see below.

  - The inner product of a bivector B ($r = 2$) and a pseudoscalar λI_3 ($s = 3$) is given by their geometric product since $\langle BI_3 \rangle_1 = BI_3$:

$$B \cdot (\lambda I_3) = \lambda B I_3 = (\lambda I_3) \cdot B \ . \tag{2.23}$$

  - Finally, the inner product of two pseudoscalars λI_3 and μI_3 ($r = s = 3$) is given by their geometric product:

$$(\lambda I_3) \cdot (\mu I_3) = -\lambda\mu = (\mu I_3) \cdot (\lambda I_3) \ . \tag{2.24}$$

- The *outer product* is indicated by a wedge. It is defined by

$$M \wedge N = \langle MN \rangle_{r+s} \tag{2.25}$$

  where M and N are two multivectors such that $M = \langle M \rangle_r$, $N = \langle N \rangle_s$, and $r + s \leq 3$. To be more specific, the following cases occur:

  - If one of M and N is a scalar ($r = 0$ or $s = 0$), the outer product is given by the scalar multiplication.
  - For two vectors ($r = s = 1$), the outer product is given according to (2.6).
  - For a vector $\mathbf{x}$ ($r = 1$) and a bivector B ($s = 2$), the outer product is given by the pseudoscalar part of the geometric product:

$$\mathbf{x} \wedge B = \langle \mathbf{x}B \rangle_3 = \frac{1}{2}(\mathbf{x}B + B\mathbf{x}) = B \wedge \mathbf{x} \ . \tag{2.26}$$

  Other outer products do not occur, due to $r + s \leq 3$. Note that in contrast to the inner product, the outer product is associative.

- Considering the various cases for the inner and outer product, the geometric product MN is always the sum of both (i.e., MN=$M \cdot N + M \wedge N$) except for the case $r = s = 2$. Given two bivectors B and C nonetheless, the geometric product BC exists. The remaining part of that geometric product, after subtracting the inner product, yields a third product, the *cross product*:

$$BC - B \cdot C = \frac{1}{2}(BC - CB) \stackrel{\text{def}}{=} B \times C = -C \times B \ . \tag{2.27}$$

Considering the last product (2.27), it is reasonable to establish a connection to the Gibbs cross product. This relation is obtained by means of *duality*. As in section 2.1, the dual of a multivector M is defined to be the product of M with the inverse of the unit pseudoscalar I_3:

$$M^* = MI_3^{-1} = -MI_3 \ . \tag{2.28}$$

By taking a closer look at figure 2.3 and table 2.2, it becomes evident that the dual of a vector $\mathbf{x}$ is the bivector which is formed by two linear independent vectors which

are orthogonal to $\mathbf{x}$. Hence, the duality maps the 3D vector space onto the 3D bivector space and vice versa. By duality, the Gibbs cross product is equivalent to the cross product of geometric algebra. The Gibbs cross product of two vectors $\mathbf{x} = x\mathbf{e}_1 + y\mathbf{e}_2 + z\mathbf{e}_3$ and $\mathbf{u} = u\mathbf{e}_1 + v\mathbf{e}_2 + w\mathbf{e}_3$ is given by

$$\mathbf{x} \times_{\mathrm{G}} \mathbf{u} = (yw - zv)\mathbf{e}_1 + (zu - xw)\mathbf{e}_2 + (xv - yu)\mathbf{e}_3 \ .$$

The cross product of $\mathbf{x}^*$ and $\mathbf{u}^*$ reads

$$(x\mathbf{e}_{23} + y\mathbf{e}_{31} + z\mathbf{e}_{12}) \times (u\mathbf{e}_{23} + v\mathbf{e}_{31} + w\mathbf{e}_{12}) = \\ (zv - yw)\mathbf{e}_{23} + (xw - zu)\mathbf{e}_{31} + (yu - xv)\mathbf{e}_{12}$$

and hence

$$(\mathbf{x} \times_{\mathrm{G}} \mathbf{u})^* = \mathbf{x}^* \times \mathbf{u}^* = \mathbf{u} \wedge \mathbf{x} \ . \tag{2.29}$$

Therefore, all three products which are known for the 3D vector space (i.e., scalar multiplication, scalar product, and Gibbs cross product) emerge from one single product, the geometric product. Having different products occurring in the algebra, some precedence rules for the products are needed: the wedge product binds strongest, and the inner product binds stronger than the geometric product.

As for the $\mathbb{R}_2$-multivector, the *modulus* of a $\mathbb{R}_3$-multivector is defined by the Euclidean norm of its components. According to (2.9), it is obtained by

$$|M| = \sqrt{\langle M\tilde{M}\rangle_0} \ , \tag{2.30}$$

where the tilde now denotes the *reversion* in $\mathbb{R}_3$. The reverse of a multivector M is defined by reversing the order of the basis elements. Hence, the scalar and vector parts are not changed, while the bivector and pseudoscalar parts are negated:

$$\tilde{M} = \langle M\rangle_0 + \langle M\rangle_1 - \langle M\rangle_2 - \langle M\rangle_3 \ . \tag{2.31}$$

As for $\mathbb{R}_2$, the reversion is an *involutive antiautomorphism* (see (2.10)). The *main automorphism* of geometric algebra is the *grade involution*:

$$\hat{M} = \langle M\rangle_0 - \langle M\rangle_1 + \langle M\rangle_2 - \langle M\rangle_3 \tag{2.32}$$

which fulfills $(MN)\hat{} = \hat{M}\hat{N}$. The combination of grade involution and reversion yields the (Clifford) *conjugation* [97]:

$$\bar{M} = \hat{\tilde{M}} = \langle M\rangle_0 - \langle M\rangle_1 - \langle M\rangle_2 + \langle M\rangle_3 \ . \tag{2.33}$$

As in the case of 2D vectors, the *inverse* of a vector $\mathbf{x}$ is given by $\mathbf{x}^{-1} = \frac{\mathbf{x}}{|\mathbf{x}|^2} = \frac{\mathbf{x}}{\mathbf{x}\cdot\mathbf{x}}$ according to (2.30). The scalar part of a geometric product as it occurs in (2.30), i.e., $\langle MN\rangle_0$, is especially simple, because the factors can be shifted periodically:

$$\langle MN\ldots PQ\rangle_0 = \langle QMN\ldots P\rangle_0 \ , \tag{2.34}$$

where $M, N, \ldots, P, Q$ are arbitrary multivectors. This identity follows from the decomposition of Q and $(MN\ldots P)$ into their components and from considering only the relevant products (i.e., those which yield a scalar result).

2.2.2 Group Representation in $\mathbb{R}_3$

Besides its interesting algebraic properties, the geometric algebra of 3D space is also capable of representing the group of rotations and isotropic scale transformations by *spinors*. By duality, these representations combine the most compact representation of a rotation, the *rotation vector*, with the numerically fastest evaluation of the group multiplication, the *quaternion* product [78].

The idea behind spinors is to generalize the interpretation of a unit complex number as a rotation. An even graded multivector (i.e., a multivector with scalar and bivector components) with unit magnitude is called a *rotor*:

$$s = a + b\mathbf{e}_{23} + c\mathbf{e}_{31} + d\mathbf{e}_{12} \tag{2.35}$$

where $a^2 + b^2 + c^2 + d^2 = 1$. As in section 2.1 an unnormalized rotor λs is called a spinor and represents a scaling-rotation. All spinors form a proper subalgebra of $\mathbb{R}_3$, which is denoted $\mathbb{R}_3^+$ and which is isomorphic to the algebra of quaternions, $\mathbb{R}_3^+ \simeq \mathbb{H}$ [97]. The action of a spinor is represented by means of a two-sided product. The reverse of the spinor s is obtained to be $\tilde{s} = a - b\mathbf{e}_{23} - c\mathbf{e}_{31} - d\mathbf{e}_{12}$, and hence in the case of a rotor $s\tilde{s} = a^2 + b^2 + c^2 + d^2 = 1$. The action of a spinor to a vector $\mathbf{x}$ is defined by $\tilde{s}\mathbf{x}s$. It is necessary to apply a spinor to a vector in this two-sided notation to avoid pseudoscalar components. In contrast to $\mathbb{R}_2$, in $\mathbb{R}_3$ there always exists a bivector which is orthogonal to a vector. If s includes a bivector orthogonal to $\mathbf{x}$, the product $\mathbf{x}s$ also results in a pseudoscalar component. This is not possible if s is applied in the two-sided way (see calculation below).

If s is a rotor, it is possible to write it as $s = \cos\varphi - \mathbf{n}^* \sin\varphi$ for an appropriate choice of φ and $\mathbf{n}$, where $\mathbf{n}$ is a unit vector. In the following the vector $\mathbf{x}$ is decomposed as $\mathbf{x}^\perp + \mathbf{x}^\parallel$, where $\mathbf{x}^\parallel = (\mathbf{x} \cdot \mathbf{n})\mathbf{n}$ is the part of $\mathbf{x}$ which is parallel to $\mathbf{n}$ (the *projection*), and $\mathbf{x}^\perp = \mathbf{x} - \mathbf{x}^\parallel = (\mathbf{x} \wedge \mathbf{n})\mathbf{n}$ is the part of $\mathbf{x}$ which is perpendicular to $\mathbf{n}$ (the *rejection*). Evaluating the action of the rotor yields

$$\begin{aligned}
\tilde{s}\mathbf{x}s &= (\cos\varphi + \mathbf{n}^*\sin\varphi)\mathbf{x}(\cos\varphi - \mathbf{n}^*\sin\varphi) \\
&= (\cos\varphi + \mathbf{n}^*\sin\varphi)(\cos\varphi - \mathbf{n}^*\sin\varphi)\mathbf{x}^\parallel + (\cos\varphi + \mathbf{n}^*\sin\varphi)^2\mathbf{x}^\perp \\
&= \mathbf{x}^\parallel + (\cos^2\varphi - \sin^2\varphi + 2\mathbf{n}^*\cos\varphi\sin\varphi)\mathbf{x}^\perp \\
&= \mathbf{x}^\parallel + (\cos(2\varphi) + \mathbf{n}^*\sin(2\varphi))\mathbf{x}^\perp \quad .
\end{aligned} \tag{2.36}$$

The term $\mathbf{n}^*\mathbf{x}^\perp$ is the vector which is orthogonal to both, $\mathbf{n}$ and $\mathbf{x}^\perp$, according to (2.29) (see also figure 2.4):

$$\mathbf{n}^*\mathbf{x}^\perp = (\mathbf{n}\mathbf{x}^\perp)^* = (\mathbf{n} \wedge \mathbf{x}^\perp)^* = \mathbf{n} \times_{\mathrm{G}} \mathbf{x}^\perp \quad .$$

Hence, $\mathbf{x}^\perp$ is changed to a linear combination of $\mathbf{x}^\perp$ and $\mathbf{n}^*\mathbf{x}^\perp$, i.e., $\mathbf{x}^\perp$ is rotated anti-clockwise around the axis defined by $\mathbf{n}$ (see figure 2.4). According to (2.36), the rotation angle is 2φ. Finally, this rotation is represented by the *rotation vector* $2\varphi\mathbf{n}$.

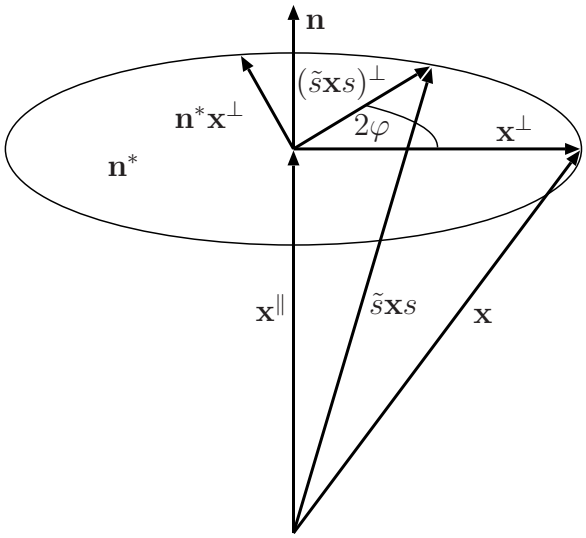

Fig. 2.4: Decomposition of a rotation in 3D space. The rotation axis is given by $\mathbf{n}$ and the rotation angle is given by 2φ. The group action to the vector $\mathbf{x}$ is $\tilde{s}\mathbf{x}s$.

In order to establish the link between the rotation vector $2\varphi\mathbf{n}$ and the normalized spinor s, the *exponential map* has to be generalized to the case of bivector arguments. The series

$$\exp B = \sum_{k=0}^{\infty} \frac{B^k}{k!} \tag{2.37}$$

converges for every bivector B, since $\exp|B|$ converges. Furthermore, the series (2.37) can be split into its scalar component and its bivector components according to $\exp B = \cos|B| + \frac{B}{|B|}\sin|B|$, which establishes a *generalized Euler equation*. The mathematical background of this equivalence is the isomorphism between the algebra of complex numbers and the subalgebra of $\mathbb{R}_3$ generated by $\{1, \frac{B}{|B|}\}$.

Applying the exponential function to the dual rotation vector (divided by two) yields

$$\exp \varphi\mathbf{n}^* = \cos\varphi + \mathbf{n}^*\sin\varphi = \tilde{s} \ , \tag{2.38}$$

i.e., the exponential function of the dual rotation vector is directly related to the normalized spinor s. On the other hand, the derivative of s with respect to the angle φ results in the unit dual rotation vector. Therefore, in terms of Lie group theory, the generator of an arbitrary 3D rotation is given by the bivector which is dual to the unit rotation vector. Interpreting the cross product (2.27) as the Lie bracket, the bivectors form a *Lie algebra*. If a specific rotation with a certain angle is considered, it is not sufficient to think of angles as scalars. In contrast to 2D, an angle in 3D has a certain directional information, and due to the representation of a rotation by the dual of its rotation vector, *angles should rather be seen as bivectors*. Hence,

a 3D angle is a bivector which is obtained from the product of the (scalar) rotation parameter and the generator. As in the 2D case, the group of scale transformations is generated by the scalars and is introduced to geometric algebra by considering general spinors.

In the 2D case, the generator of the scaling-rotation is obtained by means of the $\mathbb{R}_2^+$-logarithm (which is basically the complex logarithm). An inverse exponential map can be defined also in the 3D case. The formal definition is *identical* to that in (2.12), the only difference is that it acts on $\mathbb{R}_3^+$. Same as in the case of $\mathbb{R}_2$ where the correctness of the $\mathbb{R}_2^+$-logarithm follows from the complex logarithm, the $\mathbb{R}_3^+$-logarithm is obviously the inverse of the exponential map since both maps act on the subalgebra generated by $\{1, \frac{B}{|B|}\}$ which is isomorphic to $\mathbb{C}$.

Although it is not necessary for understanding the following chapters,[8] I prove the correctness of the logarithm by elementary calculations showing that for any bivector B, any scalar λ, and any spinor s, $\log(\exp(\lambda+B)) = \lambda+B$ and $\exp(\log(s)) = s$:

$$\begin{aligned}
\log(\exp(\lambda + B)) =& \log\left(e^{\lambda}\left(\cos|B| + \frac{B}{|B|}\sin|B|\right)\right) \\
=& \lambda + \frac{e^{\lambda}B\sin|B|}{e^{\lambda}|B||\sin|B||}\operatorname{atan}\left(\frac{e^{\lambda}|\sin|B||}{e^{\lambda}\cos|B|}\right) \\
& \text{supposing that } |B| < \pi \text{ yields} \\
=& \lambda + \frac{B}{|B|}\operatorname{atan}\left(\frac{\sin|B|}{\cos|B|}\right) = \lambda + B \ .
\end{aligned}$$

The restriction $|B| < \pi$ emerges from the periodicity of the rotation group. The rotation angle can only be recovered modulo 2π. The negative interval (i.e., $(-\pi, 0)$) is given by the orientation[9] of the plane defined by B. The special case $|B| = \pi$ is singular because the rotation axis is not well defined.

[8] However, the following proof yields deeper insight into the logarithm of $\mathbb{R}_3^+$ and also gives a hint for the implementation.

[9] The term 'orientation of a plane' from geometric algebra is a quite unlucky choice for the application to image processing where the *orientation* defines an 'unoriented' subspace whereas the *direction* defines an 'oriented' subspace. Due to this contradictory use of the term 'orientation', I prefer to use *directional sense* of a subspace instead of 'orientation of a subspace'. Hence, the difference between orientation and direction is the directional sense (see also chapter 4).

The second equality is true since

$$\begin{aligned}\exp(\log(s)) =& \exp\left(\log|s| + \frac{\langle s\rangle_2}{|\langle s\rangle_2|}\operatorname{atan}\left(\frac{|\langle s\rangle_2|}{\langle s\rangle_0}\right)\right)\\ =& |s|\left(\cos\left(\operatorname{atan}\left(\frac{|\langle s\rangle_2|}{\langle s\rangle_0}\right)\right) + \frac{\langle s\rangle_2}{|\langle s\rangle_2|}\sin\left(\operatorname{atan}\left(\frac{|\langle s\rangle_2|}{\langle s\rangle_0}\right)\right)\right)\\ & \text{since } \cos\left(\operatorname{atan}\left(\frac{|\langle s\rangle_2|}{\langle s\rangle_0}\right)\right) = \frac{\langle s\rangle_0}{|\langle s\rangle_0 + \langle s\rangle_2|} \text{ and similar for } \sin(\cdot)\\ =& |s|\left(\frac{\langle s\rangle_0}{|\langle s\rangle_0 + \langle s\rangle_2|} + \frac{\langle s\rangle_2}{|\langle s\rangle_2|}\frac{|\langle s\rangle_2|}{|\langle s\rangle_0 + \langle s\rangle_2|}\right) = s \ .\end{aligned}$$

Hence, the $\mathbb{R}_3^+$-logarithm is established.

Similar as in the 2D case, the basic results are summarized:

- $\mathbb{R}_3$ contains the vector space $\mathbb{R}^3$ including its scalar product and (Gibbs) cross product which is replaced by the wedge product. The latter represents the directed plane spanned by two vectors.
- $\mathbb{R}_3$ contains the quaternions as a subalgebra. This subalgebra of spinors represents the group action of scaling-rotations on $\mathbb{R}^3$.
- By means of duality, one can switch from spinors to rotation vectors and from planes to normal vectors.
- The geometric product is the algebra product of the 8D algebra $\mathbb{R}_3$ by which all previously mentioned products can be expressed.
  - Applied to vectors, it represents the relation of two vectors.
  - Applied to spinors, it represents the group multiplication of scaling-rotations.
  - Applied to a vector and a spinor (as a 'sandwich' product), it represents the result of the group acting on the vector.
- The three-fold outer product of vectors represents the directed volume spanned by three vectors.

Having defined the geometric algebras of 2D and 3D space, I proceed to the relationship between complex analysis and 2D harmonic vector fields as a first application of geometric algebra.

2.3 Complex Analysis and 2D Harmonic Fields

In this section parts of the commonly well known area of complex analysis is reformulated in the geometric algebra of 2D space. The also well known relationship to 2D vector fields emerges directly from the formulation in geometric algebra. In the subsequent section, the correspondence between analysis and vector fields yields a function theory for $\mathbb{R}_3$-valued functions derived from harmonic 3D vector fields which directly generalizes the idea of complex analysis.

2.3.1 The Cauchy-Riemann Equations

Complex analysis is mainly based upon the idea of holomorphic or analytic[10] functions. Analytic functions are complex functions $f(z) = f_{\mathrm{R}}(z) + i f_{\mathrm{I}}(z)$ (with $z = x + iy$)) fulfilling the *Cauchy-Riemann equations* (see e.g. [86]):

$$\begin{aligned} \frac{\partial f_{\mathrm{R}}}{\partial x} &= \frac{\partial f_{\mathrm{I}}}{\partial y} \\ \frac{\partial f_{\mathrm{R}}}{\partial y} &= -\frac{\partial f_{\mathrm{I}}}{\partial x} \end{aligned} \quad . \tag{2.39}$$

Using the isomorphism $\mathbb{R}_2^+ \simeq \mathbb{C}$ and by interpreting z as a vector instead of a scaling-rotation, z is replaced by $\mathbf{x} = \mathbf{e}_1(x + \mathbf{e}_{12}y) = x\mathbf{e}_1 + y\mathbf{e}_2$. Accordingly, the differentiation looks different:

$$\begin{aligned} \mathbf{e}_1 \frac{\partial}{\partial x} f_{\mathrm{R}} \mathbf{e}_1 &= \mathbf{e}_2 \frac{\partial}{\partial y} f_{\mathrm{I}} \mathbf{e}_2 \quad \text{and} \\ \mathbf{e}_2 \frac{\partial}{\partial y} f_{\mathrm{R}} \mathbf{e}_2 &= -\mathbf{e}_1 \frac{\partial}{\partial x} f_{\mathrm{I}} \mathbf{e}_1 \quad . \end{aligned}$$

According to the habits of geometric calculus [71], the derivative operator is always combined with the corresponding basis vector, and therefore the 2D *gradient operator*, *Dirac operator*, or *nabla operator* is defined by

$$\nabla_2 = \mathbf{e}_1 \frac{\partial}{\partial x} + \mathbf{e}_2 \frac{\partial}{\partial y} \quad . \tag{2.40}$$

Hence, the second equation of (2.39) corresponds to

$$\nabla_2 \cdot (f_{\mathrm{I}} \mathbf{e}_1 + f_{\mathrm{R}} \mathbf{e}_2) = 0 \tag{2.41}$$

[10] Analyticity means, that there is a local power series expansion at each point, which is a complete characterization of holomorphic functions [86]. For a survey of different terminologies for analytic functions see page 16 of the same reference.

and by multiplication with $\mathbf{e}_{12}$ from the right, the first one corresponds to

$$\nabla_2 \wedge (f_\mathrm{I}\mathbf{e}_1 + f_\mathrm{R}\mathbf{e}_2) = 0 \ . \tag{2.42}$$

Defining the vector field $\mathbf{f}$ by $\mathbf{f} = f_\mathrm{I}\mathbf{e}_1 + f_\mathrm{R}\mathbf{e}_2 = (f_\mathrm{R} + f_\mathrm{I}\mathbf{e}_{12})\mathbf{e}_2$, equations (2.41) and (2.42) can be combined as

$$\nabla_2 \mathbf{f} = 0 \ . \tag{2.43}$$

This equation is commonly known as the *Dirac equation* and it states that $\mathbf{f}$ has everywhere zero *divergence* and zero *curl* (rotation). Zero curl means that the vector field $\mathbf{f}$ is the *gradient field* of a potential p (short: *potential field*):

$$\mathbf{f} = \nabla_2 p \ , \tag{2.44}$$

which is easily obtained from $\nabla_2 \wedge \nabla_2 = 0$. If $\mathbf{f}$ is the gradient field of p, zero divergence means that

$$\Delta_2 p = \nabla_2 \nabla_2 p = \nabla_2 \cdot (\nabla_2 p) = \nabla_2 \cdot \mathbf{f} = 0 \ ,$$

and therefore p is a *harmonic potential* since it fulfills the *Laplace equation*

$$\Delta_2 p = \frac{\partial^2 p}{\partial x^2} + \frac{\partial^2 p}{\partial y^2} = 0 \ . \tag{2.45}$$

The gradient field of a harmonic potential is also called a *harmonic (potential) field*. The components of such a field $\mathbf{f} = f_1\mathbf{e}_1 + f_2\mathbf{e}_2$ are again harmonic:

$$\Delta_2 f_1 = \Delta_2(\mathbf{f}\cdot\mathbf{e}_1) = \frac{1}{2}(\Delta_2 \mathbf{f}\mathbf{e}_1 + \mathbf{e}_1\Delta_2\mathbf{f}) = \frac{1}{2}(\nabla_2\nabla_2\mathbf{f}\mathbf{e}_1 + \mathbf{e}_1\nabla_2\nabla_2\mathbf{f}) \overset{(2.43)}{=} \frac{1}{2}(\nabla_2 0\mathbf{e}_1 + \mathbf{e}_1\nabla_2 0) = 0$$

(similar for f_2). The harmonic scalar fields f_1 and f_2 are called the *harmonic conjugates* of each other. The relationship between harmonic functions and the solutions of the Laplace equation is explained in detail in [86].

2.3.2 Solutions of the 2D Laplace Equation

The Laplace equation itself is a partial differential equation of elliptic type [25]. The well known *fundamental solution* of the 2D Laplace equation reads

$$p_2(\mathbf{x}) = \log(|\mathbf{x}|) \ . \tag{2.46}$$

The resulting vector field is given by

$$\mathbf{f}_2(\mathbf{x}) = \nabla_2 p_2 = \frac{1}{|\mathbf{x}|}\frac{1}{2|\mathbf{x}|}2\mathbf{x} = \mathbf{x}^{-1} \tag{2.47}$$

which directly corresponds to the term z^{-1} which occurs in the Cauchy integral formula, see e.g. [26], page 353. For convenience, p_2 and $\mathbf{f}_2$ are both divided by π in order to get simpler formulas in the subsequent chapters:

$$p_2(\mathbf{x}) = \frac{\log(|\mathbf{x}|)}{\pi} \quad \text{and} \quad \mathbf{f}_2(\mathbf{x}) = (\pi\mathbf{x})^{-1} \ . \tag{2.48}$$

One fundamental theorem from complex analysis states that the *holomorphic complement* of a real function $f(x)$ on the real line exists and is unique, see [26], page 392. Extending the given real function in such a way results in a unique complex function over the complex plane which fulfills the Cauchy-Riemann equations. Hence, this extended function can also be considered as a vector field fulfilling the Dirac equation. Therefore, the holomorphic function can be identified with a harmonic field and in the light of the previous considerations, the holomorphic complement can be reduced to the following *boundary value problem of the second kind* (a *Neumann problem* [26], page 212):

$$\Delta_2 \, p(\mathbf{x}) = 0 \quad \text{for } y > 0 \text{ and} \tag{2.49}$$

$$\mathbf{e}_2 \frac{\partial}{\partial y} \, p(\mathbf{x}) = f(x)\mathbf{e}_2 \quad \text{for } y = 0 \ . \tag{2.50}$$

The first equation is just the Laplace equation restricted to the open domain $y > 0$. The second equation is called the *boundary condition* and the basis vector $\mathbf{e}_2$ is coherent with the embedding of complex functions as vector fields (the real part is embedded as the $\mathbf{e}_2$-component). The solution p of the restricted Laplace equation is made unique by the boundary condition (up to an additive constant [26], page 212). The extension of $f(x)$ to a holomorphic function is obtained from the solution of the Neumann problem by taking the gradient field of p.

In chapter 4 I will show that the Hilbert transform is given by the holomorphic complement for the limit $y \to 0$. Furthermore, I will show that y can be considered as a damping constant which introduces a *scale-space approach*. Additionally, it corresponds to the real exponent of the kernel of the *Laplace transform* [127] (see also section 3.2 in the next chapter).

2.4 3D Harmonic Fields

Applying the idea of the correspondence between complex analysis and 2D vector fields to 3D vector fields yields the generalization of complex analysis to $\mathbb{R}_3$-valued analysis.

2.4.1 The Generalized Cauchy-Riemann Equations

Starting from a 3D harmonic potential p, i.e., a solution of the 3D Laplace equation

$$\Delta_3 p = \frac{\partial^2 p}{\partial x^2} + \frac{\partial^2 p}{\partial y^2} + \frac{\partial^2 p}{\partial z^2} = 0 \ , \tag{2.51}$$

the gradient field of p is a harmonic (potential) field. From $\mathbf{f} = \nabla_3 p$, $\nabla_3 = \mathbf{e}_1 \frac{\partial}{\partial x} + \mathbf{e}_2 \frac{\partial}{\partial y} + \mathbf{e}_3 \frac{\partial}{\partial z}$, it follows that $\mathbf{f}$ fulfills the 3D Dirac equation:

$$\nabla_3 \mathbf{f} = 0 \ . \tag{2.52}$$

This equation can be split into a scalar and a bivector part:

$$\nabla_3 \cdot \mathbf{f} = 0 \tag{2.53}$$
$$\nabla_3 \wedge \mathbf{f} = 0 \tag{2.54}$$

where (2.53) states that $\mathbf{f}$ has zero divergence and (2.54) states that the curl of $\mathbf{f}$ is zero.[11] Following the same calculation as for the 2D case, it can be shown that any of the three components of $\mathbf{f}$ is again harmonic. The components form a triple of *harmonic conjugates*. They fulfill the following equalities (which are just the components of (2.52)):

$$\frac{\partial}{\partial x} f_1 + \frac{\partial}{\partial y} f_2 + \frac{\partial}{\partial z} f_3 = 0 \tag{2.55}$$
$$\frac{\partial}{\partial y} f_3 - \frac{\partial}{\partial z} f_2 = 0 \tag{2.56}$$
$$\frac{\partial}{\partial z} f_1 - \frac{\partial}{\partial x} f_3 = 0 \tag{2.57}$$
$$\frac{\partial}{\partial x} f_2 - \frac{\partial}{\partial y} f_1 = 0 \ . \tag{2.58}$$

Although I restrict myself to vector valued functions in this thesis, it should be pointed out that the Dirac equation (2.52) can also be applied to multivector-valued functions f establishing a system of generalized Cauchy-Riemann equations. A function f that fulfills (2.52) is called a *monogenic function*[12]. Using this framework of monogenic functions, most theorems from complex analysis can be generalized

[11] Applying (2.29) and (2.27) yields the classical formulation of the curl of a vector field by the Gibbs cross product.

[12] Originally, monogenic was another, somewhat archaic term for holomorphic [86]. In the Clifford analysis literature it was reused to express the multidimensional character of the functions. To be more specific, a function that fulfills (2.52) is called *left* monogenic since the Dirac operator is applied from the left. However, if $\mathbf{f}$ is a vector valued function, it is also right monogenic, and therefore I do not distinguish left and right monogenic functions in this thesis.

to higher dimensions (e.g. Cauchy's formula, residue theorem etc.). Most of these aspects, being out of the scope of this thesis, can be found in literatures on Clifford analysis [13, 98]. More recent publications on Clifford analysis also deal with PDEs and boundary value problems (see first part of [123]).

2.4.2 The 3D Laplace Equation and its Solutions

One theorem however, *is fundamental* for this thesis: the generalization of the holomorphic complement. According to the introduced notations, this generalization is called the *monogenic extension*. The monogenic extension (also: Cauchy-Kovalevska extension) of a real function $f(x,y)$ in the plane $z=0$ is the unique function which is monogenic in the upper half-space $z>0$ and which has a scalar part that is equal to $f(x,y)$ for $z=0$, see [33], page 151. In the light of the vector field representation, the monogenic extension is equivalent to a boundary value problem of the second kind:

$$\Delta_3\, p(\mathbf{x}) = 0 \quad \text{for } z>0 \text{ and} \tag{2.59}$$

$$\mathbf{e}_3 \frac{\partial}{\partial z}\, p(\mathbf{x}) = f(x,y)\mathbf{e}_3 \quad \text{for } z=0\ . \tag{2.60}$$

This boundary value problem has a unique solution p, so that $\mathbf{f}=\nabla_3\, p$ is monogenic for $z>0$, and the $\mathbf{e}_3$-component of $\mathbf{f}$ is equivalent to $f(x,y)$ for $z=0$. According to the derivation of the 1D Hilbert transform by means of the holomorphic complement (see chapter 4), the 2D generalization of the Hilbert transform will be derived from the monogenic extension. Hence, the fundamental solution of the Laplace equation and its gradient are required (which are for convenience divided by -2π):

$$p_3(\mathbf{x}) = -\frac{1}{2\pi|\mathbf{x}|} \quad \text{and} \quad \mathbf{f}_3(\mathbf{x}) = \frac{\mathbf{x}}{2\pi|\mathbf{x}|^3}\ . \tag{2.61}$$

Finally, it is important to mention that Clifford analysis is a quite new and not very common way to generalize complex analysis to higher dimensions. The better known extension is based on several complex variables [85] which is to some extent based on the n-fold tensor product of the complex algebra [31]. The drawback of using several complex variables is that the direct connection to vector field theory is lost, especially (2.55) is not obtained in that framework.

The introduced generalization of complex analysis and its relation to field theory are fundamental for understanding the 2D generalization of the analytic signal and of quadrature filters. The derivation of the generalized Hilbert transform and the related scale-space approach are based on the monogenic extension of 2D signals.

2.5 Summary of Chapter 2

In this chapter the algebraic framework which is needed for the remaining part of this thesis has been presented. Based on concepts known from complex numbers, the algebras of 2D and 3D Euclidean space have been introduced. The 2D and 3D linear vector spaces have been extended by the wedge product of vectors, which represents directed planes by means of bivectors. The subalgebras $\mathbb{R}_2^+$ and $\mathbb{R}_3^+$, which are given by the scalars and the bivectors, provide sophisticated representations of scaling-rotations in 2D and 3D Euclidean space by means of spinors. These subalgebras are isomorphic to the complex numbers and the quaternions, respectively.

The algebra product of geometric algebra is the geometric product. Applied to two vectors, it represents their geometric relation to each other, applied to two spinors, it represents the group multiplication, applied to a vector and a spinor (in 'sandwich'-notation), it represents the group action to the vector, and applied to two vectors in 'sandwich'-notation, it represents a reflection. For the algebra of 3D space, the introduction of the unit trivector, the pseudoscalar, allows to switch between spinors and rotation vectors, between planes and normal vectors, and between the Gibbs cross product and the bivector cross product by means of duality. A general trivector represents a directed volume.

Transferring the ideas of complex analysis to geometric algebra leads to the field of Clifford analysis. Straightforward generalizations of the Cauchy-Riemann equations without losing the connection to field theory are obtained. The generalization of holomorphic functions, monogenic functions, is obtained by means of the gradient field of harmonic functions, i.e., functions that fulfill the Laplace equation. The fundamental solutions of the Laplace equations in 2D and 3D are given, which establish the fundamentals to the Hilbert transform and its 2D generalization, the Riesz transform.

Chapter 3

SIGNAL THEORETIC FUNDAMENTALS

> *'He who loves practice without theory is like the sailor who boards ship without a rudder and compass and never knows where he may cast.'*
>
> Leonardo da Vinci (1452-1519)

Throughout this thesis *signals* are indicated by the symbol $\mathbf{f}$ since they are embedded as vector fields. I do this for the following reason: the Fourier transform and the Hilbert transform are closely related to complex analysis. Complex analysis, however, is related to harmonic potentials. In chapter 4 I make use of these relationships to derive a 2D generalization of the Hilbert transform. The mathematical framework I use is a boundary value problem of the second kind (i.e., with a given derivative on the boundary) where the boundary condition is given by the signal. Therefore, signals are always *vector valued*[1], whereas *filter kernels* are *spinor valued* since they act on vectors (see below). In the derivations of the Hilbert transform and its 2D generalization, I also make use of *potentials* which are *scalar valued* functions. Potentials are denoted with the lower case letter p.

The argument of a signal is a spatial vector which is represented by the notation $\mathbf{f}(\mathbf{x})$. This representation is also used for 1D signals where $\mathbf{x} = x\mathbf{e}_1$ is the 1D vector[2] with the component x. The use of 1D vectors is very important in order to be consistent to the 2D case and for the relationships to 2D vector fields and Clifford analysis. Although $\mathbf{f}(\mathbf{x})$ denotes both, 1D and 2D signals, the two cases can be distinguished by the context or by explicit remarks. All signals are supposed to behave in a tempered way, i.e., the later considered integrals do exist. Proofs about existence and convergence can be found in mathematical literature (e.g. [84, 94]) and are out of the scope of this thesis. Nevertheless, for all introduced integrals it can be shown that they exist.

[1] This is true unless multiple filter responses are combined to form the structure multivector as it is proposed at the end of chapter 5

[2] Note that 1D vector spaces are very uncommon in classical vector notation since the vector space $\mathbb{R}^1$ is normally identified with the field $\mathbb{R}$.

Linear, shift invariant (LSI) operators are multivector valued linear functionals which commute with the shift operator. In classical complex signal theory, LSI operators are identified with LSI *filters*. Due to the extended algebraic embedding in this thesis however, LSI filters are only a subset of LSI operators. Otherwise, the interpretation of the different entities (vectors, group actions, etc.) would be lost. As stated above, signals are vector valued and since filters shall be identified with operators acting on vectors, they must be *spinor valued* in order to result in a vector valued output. This would be infringed if the filter had been vector valued. LSI filters are indicated by calligraphic typeset[3]: $\mathcal{H}$. Their *impulse response* is indicated by lower case letters h and their *frequency response* by upper case letters H. Filters are not denoted with bold letters, since they are spinor valued. Note that the operator notation (calligraphic letter) of a filter is just a symbolic notation for the convolution.

3.1 The 1D Fourier Transform and LSI Operators

Introducing geometric algebra to 1D signal theory just yields some minor changes which reduce to algebraic aspects. Nevertheless, it is necessary to understand the impact of the algebraic embedding to the well known formulas.

3.1.1 Definitions

In the 1D case, the LSI filter $\mathcal{H}$ is defined by the convolution integral

$$\mathcal{H}\mathbf{f}(\mathbf{x}) = (h * \mathbf{f})(\mathbf{x}) = \int_{x'=-\infty}^{\infty} h(\mathbf{x}')\mathbf{f}(\mathbf{x}-\mathbf{x}')\,dx' \quad , \tag{3.1}$$

where the spatial vector $\mathbf{x}$ is given by $x\mathbf{e}_1$ (see also the remark about 1D vectors on page 39). The integration is performed as a classical improper Riemann integral of the first kind over the scalar valued component x' of the 1D vector $\mathbf{x}' = x'\mathbf{e}_1$ and not in the sense of a *directed integration* which would be denoted by $d\mathbf{x}$ [71]. Otherwise, a further base vector $\mathbf{e}_1$ would have been introduced and the filter output would not be vector valued any more. Furthermore, the calculation is more similar to the complex convolution, since $\mathbf{f}$ can be changed into a $\mathbb{R}_2^+$-valued function (a 'complex' function), by multiplying by $\mathbf{e}_2$ from the right: $\mathbf{f}\mathbf{e}_2 = (f_1\mathbf{e}_1 + f_2\mathbf{e}_2)\mathbf{e}_2 = f_2 + f_1\mathbf{e}_{12} \simeq f_2 + if_1$ (see also the embedding of complex functions in section 2.3). This analogy to complex convolutions is the reason for applying h from the left in (3.1). An important consequence of multiplying by $\mathbf{e}_2$ from the right is that real signals in the classical

[3] Note that besides LSI filters, the Fourier transform and the Radon transform are also indicated by calligraphic typeset, but with subsequent swift brackets, e.g. $\mathcal{F}\{\cdot\}$.

case correspond to $\mathbf{e}_2$-valued signals in the current embedding. Hence, real signals correspond to curves in the $\mathbf{e}_1 \wedge \mathbf{e}_2$-plane (see figure 3.1).

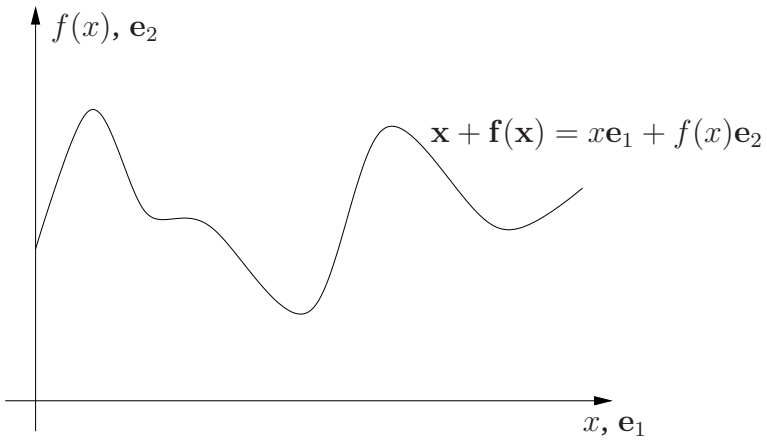

Fig. 3.1: Graph of the signal f in the $\mathbf{e}_{12}$-plane according to $\mathbf{f}(\mathbf{x}) = \mathbf{f}(x\mathbf{e}_1) = f(x)\mathbf{e}_2$.

According to the different products of geometric algebra, one can also think of different convolutions. However, those specialized convolutions do not occur in this thesis for the 1D case. Note also that the real convolution algebra (i.e., the convolution is interpreted as the algebra product and the L^2-functions form the vector space) is associative, distributive, and commutative whereas the last property is lost if the convolution algebra is extended to multivector valued functions. This is also true for the 2D case (see section 3.3).

One of the most often used LSI operators is the derivative operator. In the current embedding however, the pure derivative operator is no LSI filter by definition (3.24). According to the habits of geometric algebra and Clifford analysis, the derivative operator, or nabla operator, is a vector valued operator: $\nabla_1 g(\mathbf{x}) = \mathbf{e}_1 \frac{\partial g(\mathbf{x})}{\partial x}$ in the 1D case, and therefore it is no filter. Instead of keeping vectors, it turns odd multivector valued functions into even multivector functions and vice versa. This is especially important in the context of gradient fields which occur in the derivation of the Hilbert transform by means of vector field theory (see chapter 4). Nevertheless, it is possible to introduce an LSI filter with a similar behavior as the nabla operator by combining the latter with a vector valued Dirac delta-function. The effect of this combination is just an algebraic one: the output of the operator is again vector valued.

The relationship between the frequency response H of a filter and its impulse response h is given by the *Fourier transform*. The following notation for the 1D Fourier transform of a multivector valued function $g(\mathbf{x})$ is used throughout this thesis:

$$G(\mathbf{u}) = \mathcal{F}_1\{g(\mathbf{x})\}(\mathbf{u}) = \int_{x=-\infty}^{\infty} \exp(-\mathbf{e}_{12} 2\pi \mathbf{x}\mathbf{u}) g(\mathbf{x})\, dx \ . \tag{3.2}$$

Again, $\mathbf{x}$ and $\mathbf{u}$ are 1D vectors (i.e., $\mathbf{x} = x\mathbf{e}_1$ and $\mathbf{u} = u\mathbf{e}_1$) and their geometric product is equivalent to their inner product, $\mathbf{xu} = \mathbf{x} \cdot \mathbf{u}$. Note that the commonly used imaginary unit i is just replaced by $\mathbf{e}_{12}$, according to the isomorphism $\mathbb{R}_2^+ \simeq \mathbb{C}$. In the sequel the complex version is referred to as the *classical* case. As in the classical case, the inverse transform is given by the same formula, except for the sign of the exponent in the Fourier kernel (see e.g. [11]).

According to the Hermite symmetry of the complex Fourier transform, the transform (3.2) splits into an even part and an odd part. The even part (classical: real part) and the odd part (classical: imaginary part) of G are indicated as follows:

$$G^{\mathrm{e}}(\mathbf{u}) = \frac{G(\mathbf{u}) + G(-\mathbf{u})}{2} = \int_{x=-\infty}^{\infty} g(\mathbf{x}) \cos(2\pi \mathbf{xu})\, dx \tag{3.3}$$

$$G^{\mathrm{o}}(\mathbf{u}) = \frac{-\mathbf{e}_{12}(G(\mathbf{u}) - G(-\mathbf{u}))}{2} = -\int_{x=-\infty}^{\infty} g(\mathbf{x}) \sin(2\pi \mathbf{xu})\, dx \ . \tag{3.4}$$

Hence, it follows that $G = G^{\mathrm{e}} + \mathbf{e}_{12} G^{\mathrm{o}}$. The 1D Fourier transform is directly related to complex analysis (classical point of view) and 2D harmonic potentials (see section 2.3), yielding a direct link between harmonic oscillations and harmonic functions.

3.1.2 Basic Theorems

According to the distinction between signals, filter kernels, and potentials (being vector valued, spinor valued and scalar valued, respectively), there are three different cases of frequency domain representations. For a signal $\mathbf{f}(\mathbf{x}) = f_1(\mathbf{x})\mathbf{e}_1 + f_2(\mathbf{x})\mathbf{e}_2$, its Fourier transform $\mathbf{F}(\mathbf{u}) = \mathcal{F}_1\{\mathbf{f}(\mathbf{x})\}$ can be decomposed into its components according to

$$\mathbf{F}(\mathbf{u}) = F_1(\mathbf{u})\mathbf{e}_1 + F_2(\mathbf{u})\mathbf{e}_2 = (F_1^{\mathrm{e}}(\mathbf{u}) + F_2^{\mathrm{o}}(\mathbf{u}))\mathbf{e}_1 + (-F_1^{\mathrm{o}}(\mathbf{u}) + F_2^{\mathrm{e}}(\mathbf{u}))\mathbf{e}_2 \ , \tag{3.5}$$

where $F_1 = F_1^{\mathrm{e}} + F_1^{\mathrm{o}}\mathbf{e}_{12}$ and $F_2 = F_2^{\mathrm{e}} + F_2^{\mathrm{o}}\mathbf{e}_{12}$ are the Fourier transforms of f_1 and f_2, respectively, with their even and odd parts. Hence, the only difference to the classical complex Fourier transform of a complex signal is a subsequent multiplication with $\mathbf{e}_2$ from the right (and i is replaced by $\mathbf{e}_{12}$). The Fourier transform of a signal is denoted by a bold letter to indicate that the *spatial representation of the signal* is vector valued[4]. In the case of a spinor valued filter kernel $h = h_0 + h_{12}\mathbf{e}_{12}$, the Fourier domain representation H of h is directly given by the isomorphism $\mathbb{R}_2^+ \simeq \mathbb{C}$:

$$H(\mathbf{u}) = H_0(\mathbf{u}) + \mathbf{e}_{12} H_{12}(\mathbf{u}) = (H_0^{\mathrm{e}}(\mathbf{u}) - H_{12}^{\mathrm{o}}(\mathbf{u})) + \mathbf{e}_{12}(H_0^{\mathrm{o}}(\mathbf{u}) + H_{12}^{\mathrm{e}}(\mathbf{u})) \ , \tag{3.6}$$

which is basically the same symmetry decomposition as in the complex case. For a scalar valued potential p, its Fourier transform P is obtained from (3.6) by omitting

[4] In 1D $\mathbf{F}(\mathbf{u})$ itself is also vector valued, but this is different in 2D, see section 3.3.

the h_{12} part and replacing h_0 with p (corresponding to the complex Fourier transform of a real signal):

$$P(\mathbf{u}) = \mathcal{F}_1\{p(\mathbf{x})\}(\mathbf{u}) = P^{\mathrm{e}}(\mathbf{u}) + \mathbf{e}_{12}P^{\mathrm{o}}(\mathbf{u}) \ , \tag{3.7}$$

which is a Hermitian function in the classical case. Also in the present embedding, the conjugation can be used to express this relationship: $P(-\mathbf{u}) = \bar{P}(\mathbf{u})$.

The Fourier transform is the eigentransform of the shift operator.[5] That means that the Fourier transform is shift invariant up to a spinor factor (the linear phase, see next section). The only difference to the complex embedding is that the modulation term must be multiplied from the left:

$$\mathcal{F}_1\{g(\mathbf{x}-\mathbf{x}')\}(\mathbf{u}) = \exp(-\mathbf{e}_{12}2\pi\mathbf{x}'\mathbf{u})G(\mathbf{u}) \ , \tag{3.8}$$

where g is a general multivector valued function, and hence the shift theorem applies to both signals ($g = \mathbf{f}$) and filters ($g = h$).

Applying the shift theorem to the convolution integral (3.1) directly leads to the convolution theorem, which reads by use of the previously introduced notations

$$\mathcal{F}_1\{\mathcal{H}\mathbf{f}(\mathbf{x})\}(\mathbf{u}) = H(\mathbf{u})\mathbf{F}(\mathbf{u}) \ . \tag{3.9}$$

This is obtained since H commutes with the kernel of the Fourier transform ($H : \mathbb{R} \to \mathbb{R}_2^+$).

Strictly spoken, this convolution theorem is only valid for the application of filters to signals. If instead of a filter the derivative operator is applied to a multivector valued function g, the convolution theorem is changed a bit, yielding the derivative theorem:

$$\begin{aligned}\mathbf{e}_1\frac{\partial}{\partial x}g(\mathbf{x}) &= \mathbf{e}_1\frac{\partial}{\partial x}\mathcal{F}_1^{-1}\{G(\mathbf{u})\}(\mathbf{x}) \\ &= \mathbf{e}_1\int_{u=-\infty}^{\infty}\frac{\partial}{\partial x}\exp(\mathbf{e}_{12}2\pi\mathbf{x}\mathbf{u})G(\mathbf{u})\,du = \mathbf{e}_1\mathcal{F}_1^{-1}\{\mathbf{e}_{12}2\pi u G(\mathbf{u})\} \ ,\end{aligned}$$

and since $\mathbf{e}_1$ anti-commutes with the odd part of the Fourier kernel, this results in

$$\mathcal{F}_1\{\nabla_1 g(\mathbf{x})\} = -\mathbf{e}_1 2\pi\mathbf{u}\mathbf{e}_2 G(-\mathbf{u}) = 2\pi\mathbf{u}^* G(-\mathbf{u}) \ , \tag{3.10}$$

for any multivector valued function g.

As a consequence of (3.10), the Fourier transform of the second derivative of a function is consistent with the classical formula:

$$\begin{aligned}\mathcal{F}_1\{\nabla_1^2 g(\mathbf{x})\}(\mathbf{u}) =& 2\pi\mathbf{u}^*\mathcal{F}_1\{\nabla_1 g(\mathbf{x})\}(-\mathbf{u}) \\ =& 2\pi\mathbf{u}^*(-2\pi\mathbf{u}^* G(\mathbf{u})) = -4\pi^2\mathbf{u}^2 G(\mathbf{u})\end{aligned}$$

[5] In the discrete, finite case the shift operator can be represented as a circulant matrix with exactly one entry with value one in each column. The product of the discrete Fourier transform matrix with a circulant matrix yields a diagonal matrix. Hence, the Fourier transform matrix consists of the eigenvectors of the circulant matrix.

The operator $\nabla_1^2 = \Delta_1$ is called the 1D *Laplace operator* and according to the derivation above, it is a scalar valued operator, and hence a filter with the frequency response

$$\mathcal{F}_1\{\Delta_1\}(\mathbf{u}) = -4\pi^2\mathbf{u}^2 = -(\pm 2\pi|\mathbf{u}|)^2 \ . \tag{3.11}$$

This frequency response (and its ambiguous square root) plays a fundamental role for the derivation of the Hilbert transform (see chapter 4).

Introducing the derivative operator as a vector valued operator and therefore leaving the spinor algebra $\mathbb{R}_2^+$ (and herewith the space of filters) may appear artificial to the reader, but it is a key-point for understanding the derivation of approaches for 2D signals. *The derivatives of odd order are fundamentally different from the even order derivatives*. This fact becomes clearer in the context of vector field theory or Clifford analysis [14]. Exactly this fundamental distinction is *impossible* in the algebra of complex numbers, causing the failures in the invention of an appropriate 2D analytic signal.

The formal proofs of all mentioned theorems and further theorems (Plancherel, Parseval, uncertainty relation etc.) can be found in the elegant brief review of Fourier analysis [140] or in standard text books of signal theory [11, 114] etc., formulated in the classical framework. In the present embedding, the Parseval theorem is the same as in the complex case (just substitute the modulus of a complex number by the modulus in geometric algebra). Same is true for the uncertainty relation. I mention the result although, since it is needed for the scale-space approach in section 4.5. In 1D the uncertainty relation reads

$$(\Delta\mathbf{x})(\Delta\mathbf{u}) \geq \frac{1}{4\pi} \ , \tag{3.12}$$

which says that the product of the point spread and the frequency spread has a lower bound. In the case of a Gaussian function the equality holds [11]. The algebraic embedding has no effect to the proof of the uncertainty relation, since both involved operators (multiplication with $\mathbf{x}$ and differentiation with respect to $\mathbf{x}$) have the same base vector.

3.2 The Phase in 1D

In classical complex Fourier analysis, the spectrum of a signal or the frequency response of an operator is often considered in *amplitude and phase representation* which is simply a coordinate transform of the complex domain to polar coordinates. The amplitude spectrum is given by the modulus of the complex spectrum and the phase spectrum is obtained from the argument of the complex spectrum. Accordingly, the amplitude response and the phase response are respectively given by the modulus and the argument of the frequency response.

3.2.1 The Analytic Signal and the Local Phase

Applying the embedding into geometric algebra, the frequency space representation of a scalar function (potential)[6] or a filter is spinor valued, and therefore the amplitude response and the phase response are obtained from the logarithm (2.12):

$$|H(\mathbf{u})| = \exp(\langle \log(H(\mathbf{u})) \rangle_0) \tag{3.13}$$

$$\arg(H(\mathbf{u})) = \langle \log(H(\mathbf{u})) \rangle_2 \tag{3.14}$$

where H is the frequency response of a filter (or for the special case $\langle h \rangle_2 = 0$ of a potential). Note that the amplitude response is obtained as the exponential of the scalar part of the logarithm, i.e., the logarithm yields the sum of *attenuation* (see e.g., [129], page 210) and phase response. What is left is the amplitude-phase-approach for vector valued spectra, i.e., the Fourier representation of a signal. The idea for obtaining the latter has already been mentioned in section 2.1. A vector is simply considered as a spinor applied to the base vector $\mathbf{e}_2$.[7] The amplitude spectrum and the phase spectrum of a signal are therefore obtained by

$$|\mathbf{F}(\mathbf{u})| = \exp(\langle \log(\mathbf{F}(\mathbf{u})\mathbf{e}_2) \rangle_0) \tag{3.15}$$

$$\arg(\mathbf{F}(\mathbf{u})) = \langle \log(\mathbf{F}(\mathbf{u})\mathbf{e}_2) \rangle_2 \ . \tag{3.16}$$

The main information of a signal is contained in the phase [111], which can be used for advanced filter methods [72, 30]. Accordingly, I concentrate rather on the phase than on the amplitude in the following. By means of the phase, information is provided *what kind of structure* is contained in the signal. This information in itself is not worth too much since one needs also to know *where* this structure is located. Due to the uncertainty relation, it is not possible to obtain spatial information if the global frequency representation is considered.[8] One has to *localize* the phase and the common way to obtain a *local phase* is the *analytic signal*. By means of the analytic signal, the local structure at every point is characterized by the phase at that point (see also figure 3.2).

The analytic signal is a $\{\mathbf{e}_1, \mathbf{e}_2\}$-valued (classic: complex) signal according to

$$\mathbf{f}_{\mathrm{A}}(\mathbf{x}) = \mathbf{f}(\mathbf{x}) + \mathbf{f}_{\mathrm{H}}(\mathbf{x}) \ , \tag{3.17}$$

where $\mathbf{f}_{\mathrm{H}}$ is the *Hilbert transform* of $\mathbf{f}$, which is obtained by applying the LSI filter $\mathcal{H}_1$:

$$\mathbf{f}_{\mathrm{H}}(\mathbf{x}) = \mathcal{H}_1 \mathbf{f}(\mathbf{x}) = (h_1 * \mathbf{f})(\mathbf{x}) = \int_{x'=-\infty}^{\infty} \frac{\mathbf{e}_{12}}{\pi x'} \mathbf{f}(\mathbf{x} - \mathbf{x}')\, dx' \ . \tag{3.18}$$

[6] Since potentials will be applied as convolution kernels later, their frequency domain representation is called frequency response instead of spectrum.

[7] At this point, a fixed coordinate system is introduced.

[8] The global frequency representation has zero frequency spread and therefore, the spatial spread must be infinite.

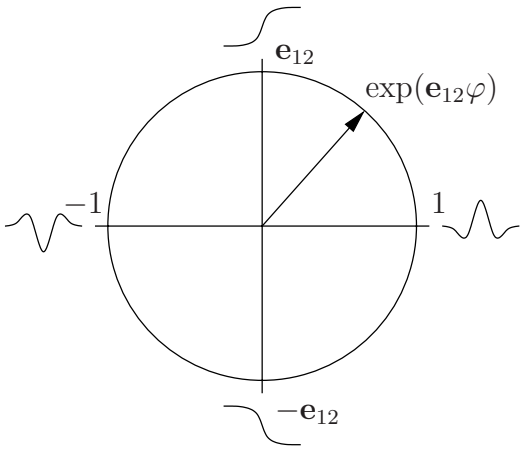

Fig. 3.2: Local phase and corresponding 1D structure. For the phases 0, $\pi/2$, π, and $-\pi/2$ the corresponding local structures are sketched. Redrawn from [61].

According to the 1D convolution theorem (3.9) and knowing that

$$\mathcal{F}_1\{(\pi x)^{-1}\}(\mathbf{u}) = -\mathbf{e}_{12}\,\mathrm{sign}(u) = \mathbf{e}_2\frac{\mathbf{u}}{|\mathbf{u}|} \quad , \tag{3.19}$$

the Fourier transform of $\mathbf{f}_\mathrm{H}$ reads

$$\mathbf{F}_\mathrm{H}(\mathbf{u}) = \mathcal{F}_1\{\mathbf{f}_\mathrm{H}(\mathbf{x})\}(\mathbf{u}) = \mathbf{e}_{12}\mathcal{F}_1\{(\pi x)^{-1}\}(\mathbf{u})\mathbf{F}(\mathbf{u}) = \mathbf{e}_1\frac{\mathbf{u}}{|\mathbf{u}|}\mathbf{F}(\mathbf{u}) \quad . \tag{3.20}$$

Consequently, the Fourier representation of $\mathbf{f}_\mathrm{A}$ is *one-sided*:

$$\mathbf{F}_\mathrm{A}(\mathbf{u}) = \mathbf{F}(\mathbf{u}) + \mathbf{F}_\mathrm{H}(\mathbf{u}) = \left(1 + \mathbf{e}_1\frac{\mathbf{u}}{|\mathbf{u}|}\right)\mathbf{F}(\mathbf{u}) = \begin{cases} 2\mathbf{F}(\mathbf{u}) & \text{if } \mathbf{u}\mathbf{e}_1 > 0 \\ \mathbf{F}(\mathbf{u}) & \text{if } \mathbf{u} = 0 \\ 0 & \text{if } \mathbf{u}\mathbf{e}_1 < 0 \end{cases} \tag{3.21}$$

Note that in (3.18) the kernel is formulated using the component x instead of using the vector $\mathbf{x}$. Actually, there is an ambiguity if (non-scalar) elements of geometric algebra occur in the denominator since the geometric product is not commutative. Fractions are commonly rewritten as the product of the enumerator and the inverse of the denominator for fixing the order, and therefore

$$\mathbf{f}_\mathrm{H}(\mathbf{x}) = \int_{x'=-\infty}^{\infty} \frac{\mathbf{e}_{12}}{\pi}\mathbf{e}_1(\mathbf{x}')^{-1}\mathbf{f}(\mathbf{x}-\mathbf{x}')\,dx' = \int_{x'=-\infty}^{\infty} \frac{\mathbf{x}'}{\pi|\mathbf{x}'|^2}\mathbf{e}_2\mathbf{f}(\mathbf{x}-\mathbf{x}')\,dx' \quad .$$

Thus, the kernel of the Hilbert transform reads in the current embedding

$$h_1(\mathbf{x}) = \frac{\mathbf{x}}{\pi|\mathbf{x}|^2}\mathbf{e}_2 \tag{3.22}$$

and the frequency response is given by

$$H_1(\mathbf{u}) = \mathbf{e}_1 \frac{\mathbf{u}}{|\mathbf{u}|} \ , \tag{3.23}$$

which means that the Hilbert transform performs a phase-shift of $\pi/2$, i.e., for a cosine the corresponding sine is obtained (see also [74]). Note that in the definition of the Hilbert transform above, the bivector which substitutes the imaginary unit does not occur in the frequency domain, but in the spatial domain.

The main difference to the classical (complex) formulation of the Hilbert transform is that the base elements of the algebra are already included in the transform.[9] According to its frequency response, two-fold application of the Hilbert transform now results in the *identity*. Furthermore, the Hilbert transform of an analytic signal results in the same analytic signal.

Self evidently, the Hilbert transform is not chosen arbitrarily and it can be derived from complex analysis, to be concrete from the Cauchy-formula [66]. Due to the equivalence of analytic or holomorphic functions and 2D harmonic potential fields, both frameworks can be used to derive the Hilbert transform. The main object of this thesis is the generalization of the Hilbert transform to two dimensions. For the Hilbert transform of 1D signals an embedding into 2D complex functions or 2D vector fields is performed. Hence, for the 2D generalization a 3D or even 4D embedding is necessary. The function theory of 3D and 4D space is less common and also quite technical, and therefore I will use vector field theory for the derivation of the 2D generalization (see chapter 4). As a motivation and in order to make the basic ideas clearer, I also derive the Hilbert transform from vector fields (see section 4.1).

3.2.2 Properties of the 1D Analytic Signal

Having defined the Hilbert transform and the analytic signal, it is reasonable to have a closer look at their properties:

1. The Hilbert transform
   - is anti-symmetric: $H_1(-\mathbf{u}) = -H_1(\mathbf{u})$.
   - suppresses the DC component: $H_1(0) = 0$.
   - is an *allpass*, i.e., its amplitude response is equal to one for all non-zero frequencies: $|H_1(\mathbf{u})| = 1 \quad \forall \mathbf{u} \neq 0$.
   - of a signal is orthogonal to the latter (see [17]).

[9] In the complex formulation (3.23) corresponds to $\mathrm{sign}(u)$ instead of $i\,\mathrm{sign}(u)$.

2. Accordingly,

   - the analytic signal of a DC free signal has twice the energy of the original signal because f and f_H are orthogonal and the Hilbert transform is an allpass.
   - the spectrum of an analytic signal is one-sided, see (3.21).
   - the analytic signal performs a *split of identity* (see below).

While some properties are irrelevant for the *characterization* of structures (e.g. the one-sidedness) others are fundamental. Especially the split of identity is of great importance. The split of identity, or its weakened form, the *invariance – equivariance* property, is obtained by considering the analytic signal in polar coordinates, i.e., in amplitude and phase representation. By applying the formulas (3.15) and (3.16), the *local amplitude* and the *local phase* of the analytic signal are obtained. The local amplitude represents the energy of a structure in the neighborhood of a certain position. Hence it varies with the local dynamic of a structure, but is invariant to changes in the structure itself. The local phase is independent of the local dynamic, but changes with the structure (see also figure 3.2). The polar coordinate system is an orthogonal coordinate system, i.e., the components are independent, and the two independent components correspond to the independent information of energy and structure. Hence, the analytic signal considered in terms of local phase and local amplitude is like an orthogonal decomposition of the information which is contained in the signal.

This is only true unless different partial signals on different *scales*, i.e., frequency bands, overlap. In this case one local phase and one local amplitude cannot represent the signal adequately, and distortions of phase and amplitude occur. To make this point clearer, suppose the signal to be decomposed into the basis functions of the Fourier transform (cosine and sine of different frequencies). Considering two of these basis functions with different frequencies and the corresponding Hilbert transforms yields the analytic signal of the interference signal (beating) which is produced by the two harmonic oscillations. Hence, the extracted local phase and amplitude are basically different from both involved local phases and amplitudes. Hence, for a general signal consisting of multiple frequency components, the decomposition of the information does not work using the analytic signal with allpass amplitude response.

In order to decompose the structure into all of its partial signals, it would be necessary to calculate the analytic signal for infinitely narrow bandwidths, i.e., Dirac-deltas in the frequency domain. According to the uncertainty relation (3.12), this would result in filters with global support, i.e., the Fourier transform (up to a modulation). However, a good approximation of the scale decomposition is given by using appropriate bandpass filters instead, in order to obtain localization in both domains,

the spatial one and the frequency one. Hence, the analytic signal is combined with a *bandpass filter* and the resulting filter set is called a *quadrature filter* (QF). Assuming sufficiently large distances between the partial signals in the frequency domain, quadrature filters with appropriate bandwidths also allow to decompose the signal into its partial signals and fulfill the invariance – equivariance property for every scale. Consequently, quadrature filters can be considered as a *compromise* between the two extrema Fourier transform and analytic signal, optimizing the uncertainty in spatial and frequency domain, see figure 3.3. Quadrature filters allow

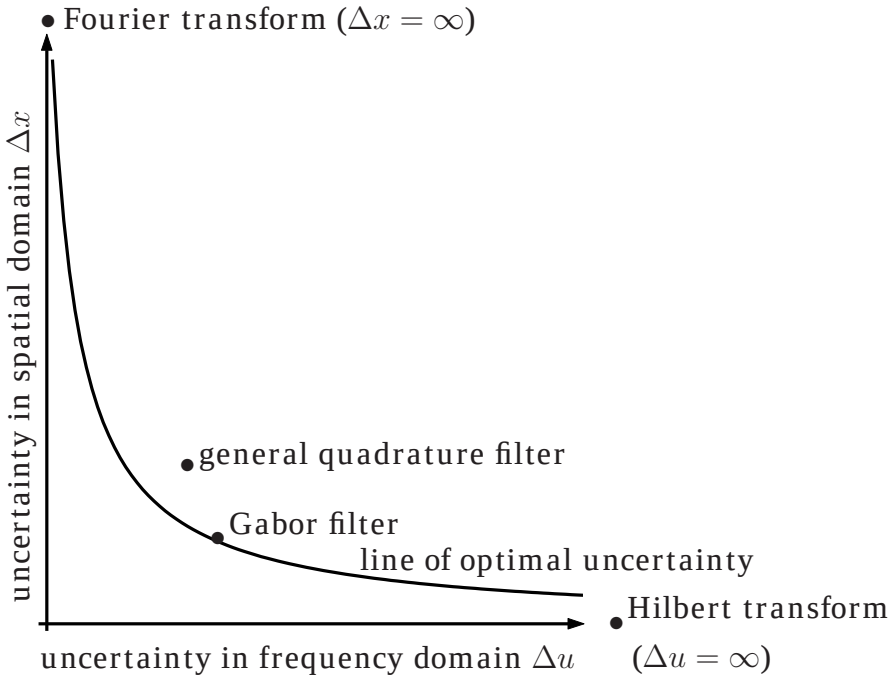

Fig. 3.3: Quadrature filters realize a compromise between infinite spatial spread (Fourier transform) and infinite frequency spread (Hilbert transform). The optimal filter with respect to the uncertainty, i.e., the Gabor filter, is not truly a quadrature filter.

to localize signals in spatial and frequency domain at the same time, i.e., the uncertainty becomes finite in both spaces. Unfortunately, it is *not possible* to construct a quadrature filter that has an uncertainty of $(4\pi)^{-1}$. One can show that the Gaussian function is the only solution for the equality in (3.12) [113], but a Gaussian function can never yield a one-sided frequency response (see (3.21)). However, it is possible to design a good approximation of a quadrature filter with uncertainty $(4\pi)^{-1}$: The *Gabor filter*. If the quotient between bandwidth and center frequency (i.e., the relative bandwidth) is small enough, the energy in the negative half-space can be kept below a certain bound. Nevertheless, the Gabor filter does not fulfill the properties of a quadrature filter in a strict sense. The problem is however, that an increasing accuracy for the phase estimation yields an increasing number of Gabor filters

to cover a certain passband. If the aim is an accurate phase estimation with low computational load, a filter with suboptimal uncertainty is certainly a better choice than the Gabor filter. This problem is even worse in 2D, since decreasing relative bandwidth does not only mean more Gabor filters to cover a certain passband, but also more samples of the local orientation.[10]

3.3 The 2D Fourier Transform and Linear Operators

According to the embedding of real 1D signals into the $\mathbf{e}_1 \wedge \mathbf{e}_2$-plane (see figure 3.1), real 2D signals (e.g. images) are considered as *surfaces* in 3D space (see figure 3.4 and [76], page 163). In the case of images, it is important to notice that the 2D

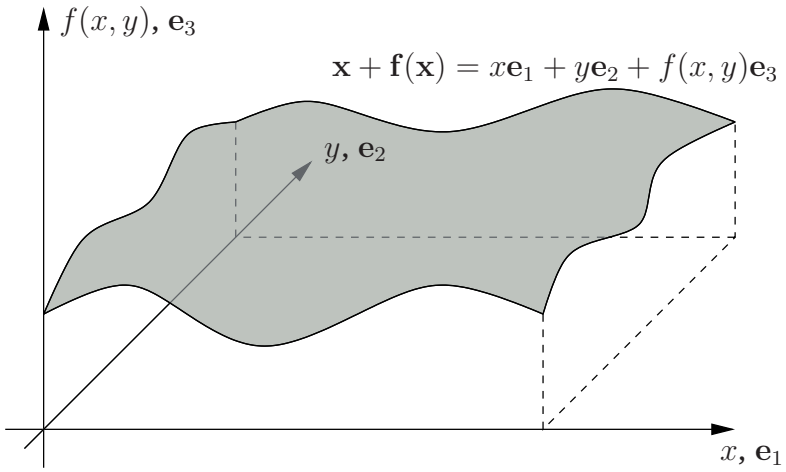

Fig. 3.4: Graph of the 2D signal f in 3D space according to $\mathbf{f}(\mathbf{x}) = \mathbf{f}(x\mathbf{e}_1 + y\mathbf{e}_2) = f(x,y)\mathbf{e}_3$.

signal $f(x,y)$ can be any representation of a grayscale image which emerges from a strict monotonic function. In [81], for example, the logarithm of the intensity is used (although the third base vector is supposed to be a *null vector* in that paper which yields some problems defining a harmonic field in that space).

3.3.1 Definitions

In the light of the definition of a 1D LSI filter and using the representation of group actions in $\mathbb{R}_3$ (see (2.36)), a 2D LSI filter should be defined as an integral

[10] The 2D filter with optimal uncertainty is not a general Gaussian function, but an isotropic Gaussian function, see next section.

of the 'sandwich' product $\tilde{h}'\mathbf{f}h'$. For special choices of $\mathbf{f}$ and h', $\mathbf{f}$ and $\langle h' \rangle_2$ are anti-commutative. Hence, there exists a spinor $h = \tilde{h}'\tilde{h}'$ so that $\tilde{h}'\mathbf{f}h' = \tilde{h}'\tilde{h}'\mathbf{f} = h\mathbf{f}$, which means that both formulations, the one-sided and the sandwich product, are equivalent. The case of anti-commuting operator and signal occurs in all derivations for the intrinsically 1D case (see chapter 4), but for the i2D case it becomes necessary to use also commuting operators. If $\mathbf{f}$ and $\langle h' \rangle_2$ are not anti-commutative, the result of $\tilde{h}'\mathbf{f}h'$ differs from $h\mathbf{f}$ in general. For the interpretation of the 2D convolution and for the convolution theorem however, the latter formulation is preferable, and therefore a 2D LSI filter $\mathcal{H}$ is defined by

$$\mathcal{H}\mathbf{f}(\mathbf{x}) = (h * \mathbf{f})(\mathbf{x}) = \int_{y'=-\infty}^{\infty} \int_{x'=-\infty}^{\infty} h(\mathbf{x}')\mathbf{f}(\mathbf{x} - \mathbf{x}')\, dx'\, dy' \quad , \tag{3.24}$$

where $h(\mathbf{x})$ is $\mathbb{R}_3^+$-valued. Note that the filter output is not vector-valued in general due to the multiplication from the left instead of the two-sided spinor notation. The question whether there is an algebraic embedding in a higher dimensional space where the sandwich product can be used for the convolution remains open; at least the simple extensions to algebras of 4D and 5D Euclidean space are not suffiecient to realize such an embedding.

The one-sided notation results in a simpler convolution theorem (see (3.32)) and what is even more striking, it is appropriate for the application of commutative *and* anti-commutative operators. For the anti-commutative operators (i.e., $\langle h(\mathbf{x}) \rangle_2 \in \text{span}\{\mathbf{e}_{23}, \mathbf{e}_{31}\}$ if $\mathbf{f}(\mathbf{x})$ is $\mathbf{e}_3$-valued) the one-sided product is identical to the two-sided product (as explained above) and can therefore be considered as a geometric operation. For the commutative operators (i.e., $\langle h(\mathbf{x}) \rangle_2$ is $\mathbf{e}_{12}$-valued), the two-sided product mixes scalar and trivector components, in contrast to the one-sided notation which rather corresponds to a complex multiplication. Hence, the commutative operators *directly correspond to complex valued operators in the classical framework*, whereas the anti-commutative operators *introduce a new class of linear operators* which cannot be represented in an adequate way in the algebra of complex numbers.

As in the 1D case, the derivative operator $\nabla_2 g(\mathbf{x}) = \mathbf{e}_1 \frac{\partial g(\mathbf{x})}{\partial x} + \mathbf{e}_2 \frac{\partial g(\mathbf{x})}{\partial y}$ is not a filter according to definition (3.24). Again, the derivative operator must be combined with a vector valued Dirac delta-function in order to apply the methods which are derived for the LSI filters in this section.

In order to obtain the frequency response of a 2D LSI filter, the 2D Fourier transform must be applied. Basically, the 2D Fourier transform is the concatenation of 1D Fourier transforms in x- and y-direction, which is equal to the 2D integral transform with the kernel $\exp(-i2\pi\mathbf{x} \cdot \mathbf{u})$ in the complex framework [12]. At this point, it is very important to stress the fact that the scalar product is invariant under rotations of the vector space (i.e., rotations are isometries [24]). Therefore, the 2D Fourier transform relates a Cartesian approach (concatenation of 1D Fourier trans-

forms) with a polar point of view since it is *isotropic*. The Fourier transform does not depend on the choice of the (orthogonal) basis vectors of the plane. More generally, an isotropic operator commutes with the group of rotations.

The algebraically embedded Fourier transform of a 2D function is denoted by a capital letter, as in the 1D case, and it is defined by:

$$G(\mathbf{u}) = \int_{y=-\infty}^{\infty} \int_{x=-\infty}^{\infty} g(\mathbf{x}) \exp(-I_3 2\pi \mathbf{x} \cdot \mathbf{u})\, dx\, dy \quad . \tag{3.25}$$

This definition yields an isotropic transform, like in the classical complex case and in contrast to the quaternionic Fourier transform, see [21]. Note that in the 2D case the imaginary unit is not replaced by $\mathbf{e}_{12}$, but by $I_3 = \mathbf{e}_{123}$. Since I_3 commutes with every element of the algebra, the transform kernel can be placed on the left or on the right, whereas in the 1D case it has to be on the left. Just as $\mathbb{R}_2^+$, the algebra of $\{1, I_3\}$ is isomorphic to the complex numbers [97], page 55. Multiplying an element of this algebra with a unit vector $\mathbf{n}$ yields an element in $\{\mathbf{n}, -\mathbf{n}^*\}$, which can be identified with an element of a complex vector space. In the context of the Fourier transform, I consider three subspaces of this complex vector space: $\{\mathbf{e}_1, \mathbf{e}_{23}\}$, $\{\mathbf{e}_2, \mathbf{e}_{31}\}$, and $\{\mathbf{e}_3, \mathbf{e}_{12}\}$.

According to (3.3) and (3.4), the 2D transform can also be split into even and odd parts:

$$G^{\mathrm{e}}(\mathbf{u}) = \frac{G(\mathbf{u}) + G(-\mathbf{u})}{2} = \int_{y=-\infty}^{\infty} \int_{x=-\infty}^{\infty} g(\mathbf{x}) \cos(2\pi \mathbf{x} \cdot \mathbf{u})\, dx\, dy \tag{3.26}$$

$$G^{\mathrm{o}}(\mathbf{u}) = \frac{(G(\mathbf{u}) - G(-\mathbf{u}))^*}{2} = -\int_{y=-\infty}^{\infty} \int_{x=-\infty}^{\infty} g(\mathbf{x}) \sin(2\pi \mathbf{x} \cdot \mathbf{u})\, dx\, dy \tag{3.27}$$

and it follows that $G = G^{\mathrm{e}} + G^{\mathrm{o}} I_3$.

Note that the definition of even and odd functions is *the fundamental design principle* for the whole following signal theory. While the definition of symmetry in 1D is unique, there are at least two principles of symmetry in 2D: line symmetry and point symmetry. I discuss this topic in detail in chapter 4, but there are some fundamental ideas which are already important in this section. Choosing line symmetry might appear more obvious then using point symmetry and indeed, there are some approaches to this kind of generalized Fourier transform (e.g. the quaternionic Fourier transform [38, 37, 22], or more general the Clifford Fourier transform [13, 21], or Fourier transforms on complex tensor spaces [40, 42]). In the last reference, it is shown that all these transforms are basically based on a tensor approach. The resulting embedding space is 4D, reflecting the four possible line symmetries of 2D signals in a rectangular coordinate system. Without going into technical details, the Fourier spectrum is obtained as a 2D subspace of a 4D generalized Laplace

transform domain, consisting of two frequency variables and two distinct decay parameters attached to the axes. Each of the damping parameters affects that part of the basis functions (hypercomplex exponentials) that varies along its corresponding coordinate axis. Consequently, any embedding using a 4D space depends on the specific choice of the coordinate system, and hence none of these approaches can be isotropic.

In my approach however, I use an embedding into 3D space, motivated by point symmetries and isotropy. As far as I know, there is no approach in signal theory or image processing that relates harmonic 3D functions with the 2D Fourier transform. However, the mathematicians developed such a theory [94], but unfortunately their results are not directly applicable to the problem of a 2D analytic signal. Exactly this is the main contribution described in chapter 4. In my approach the third coordinate of the embedding space is a decay parameter which is common for both frequency coordinates.

Summarizing this short discussion, one can conclude that the isotropic Fourier transform according to (3.25) is *inherently* related to point symmetry while the quaternionic Fourier transform is related to line symmetries.

3.3.2 Basic Theorems

Returning to the 2D Fourier transform itself, several theorems are now established corresponding to those in section 3.1. At first, the Fourier transforms of signals, filters, and potentials are considered more closely. The Fourier domain representation of a signal $\mathbf{f}(\mathbf{x}) = f_1(\mathbf{x})\mathbf{e}_1 + f_2(\mathbf{x})\mathbf{e}_2 + f_3(\mathbf{x})\mathbf{e}_3$ decomposes as follows:

$$\begin{aligned}\mathbf{F}(\mathbf{u}) &= \mathcal{F}_2\{\mathbf{f}(\mathbf{x})\}(\mathbf{u}) = \mathbf{e}_1 F_1(\mathbf{u}) + \mathbf{e}_2 F_2(\mathbf{u}) + \mathbf{e}_3 F_3(\mathbf{u}) \\ &= \mathbf{e}_1 F_1^{\mathrm{e}}(\mathbf{u}) - \mathbf{e}_1^* F_1^{\mathrm{o}}(\mathbf{u}) + \mathbf{e}_2 F_2^{\mathrm{e}}(\mathbf{u}) - \mathbf{e}_2^* F_2^{\mathrm{o}}(\mathbf{u}) + \mathbf{e}_3 F_3^{\mathrm{e}}(\mathbf{u}) - \mathbf{e}_3^* F_3^{\mathrm{o}}(\mathbf{u}) \ ,\end{aligned} \tag{3.28}$$

where $F_k = F_k^{\mathrm{e}} + F_k^{\mathrm{o}} I_3$ is the Fourier transforms of f_k with its even and odd part for $k = 1, 2, 3$. Hence, the fundamental difference to the 1D Fourier transform (3.2) and to the complex 2D signal theory is that vector valued signals are transformed such that the three components are kept separated. Note that there is no I_3-component in the spectral representation, hence there is no component which directly corresponds to *the* complex component according to the isomorphism $i \simeq I_3$. Instead, each of the three components of the vector field has its own bivector (replacing the imaginary part) in the frequency domain. Although $\mathbf{F}(\mathbf{u})$ is not vector valued in general, it is denoted by a bold face letter, indicating that it is the Fourier representation of a vector valued signal (see also footnote 4 on page 42). The symmetry relation in (3.28) can be summarized by $\mathbf{F}(-\mathbf{u}) = \tilde{\mathbf{F}}(\mathbf{u})$, i.e., the spectrum of a signal is symmetric with respect to reversion.

For a (spinor valued) filter $h = h_0 + h_{23}\mathbf{e}_{23} + h_{31}\mathbf{e}_{31} + h_{12}\mathbf{e}_{12}$, the frequency response decomposes as follows:

$$\begin{aligned} H(\mathbf{u}) &= \mathcal{F}_2\{h(\mathbf{x})\}(\mathbf{u}) \\ &= H_0(\mathbf{u}) + \mathbf{e}_{23}H_{23}(\mathbf{u}) + \mathbf{e}_{31}H_{31}(\mathbf{u}) + \mathbf{e}_{12}H_{12}(\mathbf{u}) \\ &= H_0^{\mathrm{e}}(\mathbf{u}) + H_0^{\mathrm{o}}(\mathbf{u})I_3 + \mathbf{e}_{23}H_{23}^{\mathrm{e}}(\mathbf{u}) + \mathbf{e}_1 H_{23}^{\mathrm{o}}(\mathbf{u}) \\ &\quad + \mathbf{e}_{31}H_{31}^{\mathrm{e}}(\mathbf{u}) + \mathbf{e}_2 H_{31}^{\mathrm{o}}(\mathbf{u}) + \mathbf{e}_{12}H_{12}^{\mathrm{e}}(\mathbf{u}) + \mathbf{e}_3 H_{12}^{\mathrm{o}}(\mathbf{u}) \ . \end{aligned} \tag{3.29}$$

Hence, all components of the filter are transformed into distinct components (where $H_0 = \mathcal{F}_2\{h_0\}$, $H_{23} = \mathcal{F}_2\{h_{23}\}$ and so on). The symmetry relation in (3.29) can be summarized as $H(-\mathbf{u}) = \hat{H}(\mathbf{u})$, i.e., the frequency response of a filter is symmetric with respect to grade involution.

As a special case of the spinor valued function, the Fourier domain representation of a potential p is obtained by

$$P(\mathbf{u}) = \mathcal{F}_2\{p(\mathbf{x})\}(\mathbf{u}) = P^{\mathrm{e}}(\mathbf{u}) + P^{\mathrm{o}}(\mathbf{u})I_3 \ , \tag{3.30}$$

which fulfills both symmetry relations at a time ($P(-\mathbf{u}) = \tilde{P}(\mathbf{u}) = \hat{P}(\mathbf{u})$).

Just as the 1D Fourier transform, the 2D Fourier transform is shift invariant:

$$\mathcal{F}_2\{g(\mathbf{x}-\mathbf{x}')\}(\mathbf{u}) = G(\mathbf{u})\exp(-I_3 2\pi\mathbf{x}'\cdot\mathbf{u}) = \exp(-I_3 2\pi\mathbf{x}'\cdot\mathbf{u})G(\mathbf{u}) \ , \tag{3.31}$$

where g is a general multivector valued function.

Applying the 2D shift theorem to the 2D convolution (3.24) directly leads to the convolution theorem, which reads by use of the previously introduced notations

$$\mathcal{F}_2\{\mathcal{H}\mathbf{f}(\mathbf{x})\}(\mathbf{u}) = H(\mathbf{u})\mathbf{F}(\mathbf{u}) \ . \tag{3.32}$$

This is obtained from the fact that the 2D Fourier kernel commutes with any multivector and especially with H. In the context of (3.24), I mentioned that a pseudoscalar part can occur in the operator output. Due to the definition of the 2D Fourier transform, the pseudoscalar part in the filter output corresponds to an even pseudoscalar part and an odd scalar part in the spectrum. The (even) pseudoscalar part reads $\langle H\rangle_1 \wedge \langle\mathbf{F}\rangle_2 + \langle H\rangle_2 \wedge \langle\mathbf{F}\rangle_1$, whereas the (odd) scalar part reads $\langle H\rangle_1 \cdot \langle\mathbf{F}\rangle_1 + \langle H\rangle_2 \cdot \langle\mathbf{F}\rangle_2$. Combining both parts gives $\frac{1}{2}(H\mathbf{F} - \mathbf{F}\bar{H})$ which is the commutator of the frequency response of $\langle h\rangle_2$ and of $\mathbf{F}$. A commutative filter in $\mathbb{R}_3^+$ consists of this commutator part and a real part in general. This class of filters directly corresponds to complex filters (i.e., the filters which can be applied in the classical framework). The non-commuting filters establish a *new class of filters* which can only be applied in the framework of geometric algebra. In this thesis I concentrate on the latter, and therefore the filter output is vector valued except for some special cases which are explicitly indicated.

In 2D the convolution theorem does not only apply to LSI filters, but also to the more general class of LSI operators, since the Fourier kernel commutes with any multivector in $\mathbb{R}_3$. A special case for a not spinor-valued LSI operator is again the derivative operator. The Fourier transform of the gradient of a multivector valued function g is obtained in a similar way as for the 1D case:

$$\begin{aligned}\nabla_2 g(\mathbf{x}) &= (\mathbf{e}_1\frac{\partial}{\partial x} + \mathbf{e}_2\frac{\partial}{\partial y})\mathcal{F}_2^{-1}\{G(\mathbf{u})\}(\mathbf{x}) \\ &= \int_{v=-\infty}^{\infty}\int_{u=-\infty}^{\infty}(\mathbf{e}_1\frac{\partial}{\partial x} + \mathbf{e}_2\frac{\partial}{\partial y})G(\mathbf{u})\exp(I_3 2\pi(\mathbf{x}\cdot\mathbf{u}))\,du\,dv \\ &= \int_{v=-\infty}^{\infty}\int_{u=-\infty}^{\infty}2\pi(u\mathbf{e}_1 + v\mathbf{e}_2)I_3 G(\mathbf{u})\,du\,dv \ ,\end{aligned}$$

and therefore

$$\mathcal{F}_2\{\nabla_2 g(\mathbf{x})\}(\mathbf{u}) = -2\pi\mathbf{u}^* G(\mathbf{u}) \ . \tag{3.33}$$

As in the 1D case, two-fold application of the derivative theorem results in the frequency response of the Laplace operator:

$$\mathcal{F}_2\{\Delta_2\}(\mathbf{u}) = -4\pi^2\mathbf{u}^2 = -(\pm 2\pi|\mathbf{u}|)^2 \ , \tag{3.34}$$

which is again an LSI filter.

A further fundamental theorem of 2D Fourier analysis is the affine theorem [12]. In contrast to the previous theorems, the dimensionality of the embedding space yields a *new quality* of relationships for the affine group. At first, the transformation group is restricted to scale transformations and rotations in the $\mathbf{e}_{12}$-plane. According to section 2.2, a scaling-rotation is represented as the spinor $s = \sqrt{\lambda}\exp(\mathbf{e}_{12}\theta/2)$ ($\lambda > 0$). Hence, the affine theorem reads

$$\mathcal{F}_2\{g(\tilde{s}\mathbf{x}s)\}(\mathbf{u}) = \lambda^{-2}G(\lambda^{-2}\tilde{s}\mathbf{u}s) = \lambda^{-2}G(s^{-1}\mathbf{u}\tilde{s}^{-1}) \ . \tag{3.35}$$

However, in chapter 5 the affine theorem is needed in a more general setting. The considered group of transformations is extended by general dilations, i.e., the scale is altered differently in two directions in the $\mathbf{e}_{12}$-plane. In general these group actions cannot be represented by a spinor in the current embedding. It is necessary to consider *projections* of the vector $\mathbf{x}$ onto the main axes $\mathbf{n}$ and $\mathbf{n}I_2^{-1} = -\mathbf{n}I_2$ in the $\mathbf{e}_{12}$-plane. The projection onto $\mathbf{n}$ is obtained by $(\mathbf{n}\cdot\mathbf{x})\mathbf{n}$ whereas the projection onto $-\mathbf{n}I_2$ (the orthogonal vector) is given by the *rejection*: $((\mathbf{n}I_2)\cdot\mathbf{x})(\mathbf{n}I_2) = \mathbf{n}(\mathbf{n}\wedge\mathbf{x})$ (for details of this calculation see (5.6) in chapter 5). Using the projections, the group of anisotropic dilations in the $\mathbf{e}_{12}$-plane can be represented as

$$\mathbf{x}' = \lambda_1\mathbf{n}(\mathbf{n}\cdot\mathbf{x}) + \lambda_2\mathbf{n}(\mathbf{n}\wedge\mathbf{x}) \ , \tag{3.36}$$

with $\lambda_1, \lambda_2 \in \mathbb{R}^+$ and $\mathbf{n}$ being a unit vector in the $\mathbf{e}_{12}$-plane. Accordingly,

$$\mathbf{x} = \lambda_1^{-1}\mathbf{n}(\mathbf{n}\cdot\mathbf{x}') + \lambda_2^{-1}\mathbf{n}(\mathbf{n}\wedge\mathbf{x}') \tag{3.37}$$

yields the inverse transformation. Hence, the 2D Fourier transform of a general function $g(\mathbf{x}')$ is obtained to be

$$\begin{aligned}\mathcal{F}_2\{g(\mathbf{x}')\} &= \iint g(\mathbf{x}') \exp(-I_3 2\pi \mathbf{u}\cdot\mathbf{x})\, dx\, dy \\ &= \iint g(\mathbf{x}') \exp(-I_3\pi \mathbf{u}\cdot(\lambda_1^{-1}\mathbf{n}(\mathbf{n}\cdot\mathbf{x}') + \lambda_2^{-1}\mathbf{n}(\mathbf{n}\wedge\mathbf{x}')))\, \frac{dx'\, dy'}{\lambda_1\lambda_2} \\ &= \frac{1}{\lambda_1\lambda_2}\iint g(\mathbf{x}') \exp(-I_3\pi\mathbf{u}'\cdot\mathbf{x}')\, dx'\, dy' = \frac{1}{\lambda_1\lambda_2} G(\mathbf{u}') \ , \end{aligned} \tag{3.38}$$

where $\mathbf{u}'$ is defined by $\mathbf{u}'\cdot\mathbf{x}' = \mathbf{u}\cdot(\lambda_1^{-1}\mathbf{n}(\mathbf{n}\cdot\mathbf{x}') + \lambda_2^{-1}\mathbf{n}(\mathbf{n}\wedge\mathbf{x}'))$ which yields

$$\mathbf{u}' = \lambda_1^{-1}\mathbf{n}(\mathbf{n}\cdot\mathbf{u}) + \lambda_2^{-1}\mathbf{n}(\mathbf{n}\wedge\mathbf{u}) \ .$$

3.3.3 Plancherel Theorem and 2D Uncertainty

In the subsequent paragraphs, I prove the Plancherel theorem (see e.g. [140] for the embedded signals and as a consequence the Parseval theorem is derived. To understand the idea of the Plancherel theorem one has to switch from the geometric algebras of 2D or 3D space to the infinite dimensional vector space of functions. In the classical case, the Plancherel theorem is formulated for the *complex* function space. However in the present case, the algebra to consider is $\mathbb{R}_3$ and the scalar product in that vector space is defined by

$$\langle g_1, g_2\rangle = \int_{y=-\infty}^{\infty}\int_{x=-\infty}^{\infty} g_1(\mathbf{x})\tilde{g}_2(\mathbf{x})\, dx\, dy \tag{3.39}$$

where g_1 and g_2 are arbitrary multivector valued functions. Hence this scalar product is multivector valued in general. By means of this scalar product, the Plancherel theorem reads

$$\langle g_1, g_2\rangle = \langle G_1, G_2\rangle \tag{3.40}$$

where $G_1 = \mathcal{F}_2\{g_1\}$ and $G_2 = \mathcal{F}_2\{g_2\}$. The theorem is proved by application of the 2D convolution theorem (all double integrals are over the whole space $\mathbb{R}_2$):

$$\begin{aligned}\langle g_1, g_2\rangle &= \iint g_1(\mathbf{x})\tilde{g}_2(\mathbf{x})\, dx\, dy \\ &= \iint\iint G_1(\mathbf{u}) \exp(I_3 2\pi\mathbf{u}\cdot\mathbf{x})\, du\, dv \iint \tilde{G}_2(\mathbf{u}') \exp(-I_3 2\pi\mathbf{u}'\cdot\mathbf{x})\, du'\, dv'\, dx\, dy \\ &= \iint\iint G_1(\mathbf{u})\tilde{G}_2(\mathbf{u}') \iint \exp(I_3 2\pi(\mathbf{u}-\mathbf{u}')\cdot\mathbf{x})\, dx\, dy\, du\, dv\, du'\, dv' \\ &= \iint\iint G_1(\mathbf{u})\tilde{G}_2(\mathbf{u}')\delta_0(\mathbf{u}-\mathbf{u}')\, du'\, dv'\, du\, dv \\ &= \iint G_1(\mathbf{u})\tilde{G}_2(\mathbf{u})\, du\, dv = \langle G_1, G_2\rangle \end{aligned}$$

As a consequence of (3.40), each component of the scalar product is equal:

$$\langle\langle g_1, g_2\rangle\rangle_k = \langle\langle G_1, G_2\rangle\rangle_k \tag{3.41}$$

for any $k \in \{0, 1, 2, 3\}$. For $k = 0$ and $g = g_1 = g_2$ the Parseval theorem is obtained:

$$E_g = \langle\langle g, g\rangle\rangle_0 = \langle\langle G, G\rangle\rangle_0 = E_G \tag{3.42}$$

where E_g denotes the energy of g and E_G denotes the energy of G.

A fundamental theorem about the 1D Fourier transform is the uncertainty relation, which can be proven simply by using the Plancherel theorem [140]. In 2D however, the uncertainty relation is still an open problem. In [61] it is stated that there is 'no straightforward formulation' for the 2D uncertainty relation. According to [120] (page 221) the 2D uncertainty relation is formulated as

$$\left(\iint |\mathbf{x}|^2 |g(\mathbf{x})|^2 \, dx \, dy\right)^{1/2} \left(\iint |\mathbf{u}|^2 |G(\mathbf{u})|^2 \, du \, dv\right)^{1/2} \geq \frac{1}{2\pi} E_g \ , \tag{3.43}$$

which can be abbreviated as

$$(\Delta\mathbf{x})(\Delta\mathbf{u}) \geq \frac{1}{2\pi} \ . \tag{3.44}$$

This says that the product of the *modulus* of the point spread and the *modulus* of the frequency spread has a lower bound. Omitting the directional information, the author circumvents a fundamental problem of the 2D case. The latter is hidden behind the definition of the variance. In the equation above, solely the *distance* to the origin is taken into account. Hence, the resulting solution is *cumulated* in the sense that all components are added:

$$\sigma_{\mathbf{x}}^2 \sigma_{\mathbf{u}}^2 = (\sigma_x^2 + \sigma_y^2)(\sigma_u^2 + \sigma_v^2) = (\sigma_x^2 \sigma_u^2 + \sigma_y^2 \sigma_v^2) + (\sigma_x^2 \sigma_v^2 + \sigma_y^2 \sigma_u^2) \tag{3.45}$$

In the case of separable signals, the cumulated equation (3.45) can be split into two uncertainty relations. The first expression $\sigma_x^2 \sigma_u^2 + \sigma_y^2 \sigma_v^2$ is just the sum of the variances of the 1D signals in which the 2D signal can be separated. The uncertainty of the remaining term in (3.45) is given as

$$\sigma_x^2 \sigma_v^2 + \sigma_y^2 \sigma_u^2 \geq \frac{1}{4\pi^2} - \sigma_x^2 \sigma_u^2 - \sigma_y^2 \sigma_v^2 \ . \tag{3.46}$$

Unfortunately, it is not possible to give a general lower bound for the expression $\sigma_x^2 \sigma_v^2 + \sigma_y^2 \sigma_u^2$ unless the uncertainties of the two involved 1D signals is calculated.

For non-separable signals it is even worse since the uncertainty relation can only be stated in its cumulated form (3.45). Nevertheless it is reasonable to split the uncertainty into its two parts for specific cases. The two parts represent the parallel uncertainty and the orthogonal uncertainty. The latter can be considered as

a measure for the isotropy of the signal and can be embedded as an $\mathbf{e}_{12}$-component. Both uncertainties can be combined as

$$(\Delta\mathbf{x})(\Delta\mathbf{u}) = ((\Delta\mathbf{x})(\Delta\mathbf{u}))^{\|} + ((\Delta\mathbf{x})(\Delta\mathbf{u}))^{\perp}\mathbf{e}_{12} \tag{3.47}$$

so that

$$|(\Delta\mathbf{x})(\Delta\mathbf{u})| = \sqrt{((\Delta\mathbf{x})(\Delta\mathbf{u}))^{\|2} + ((\Delta\mathbf{x})(\Delta\mathbf{u}))^{\perp 2}} \geq (2\pi)^{-1}$$

according to (3.44).

Consider, e.g., the 2D Gaussian hump with different standard deviations σ_1 and σ_2 in x- and y-direction. The uncertainty along either of the axes is $(4\pi)^{-1}$ which follows from the 1D uncertainty principle. Accordingly, $((\Delta\mathbf{x})(\Delta\mathbf{u}))^{\|} = \sqrt{2}(4\pi)^{-1}$, and therefore the mixed terms (3.46) can be estimated by $((\Delta\mathbf{x})(\Delta\mathbf{u}))^{\perp} \geq \sqrt{2}(4\pi)^{-1}$. Performing the calculations with $\sigma_1 \neq \sigma_2$ yields

$$((\Delta\mathbf{x})(\Delta\mathbf{u}))^{\perp} = \frac{1}{4\pi}\sqrt{\frac{\sigma_1^2}{\sigma_2^2} + \frac{\sigma_2^2}{\sigma_1^2}} \ .$$

For $\sigma_1 = \sigma_2$ this results in $((\Delta\mathbf{x})(\Delta\mathbf{u}))^{\perp} = \sqrt{2}(4\pi)^{-1}$ which is the optimal case. For anisotropic Gaussian humps the uncertainty is suboptimal and the orthogonal part is proportional to

$$\sqrt{\frac{\sigma_1^2}{\sigma_2^2} + \frac{\sigma_2^2}{\sigma_1^2}} = \frac{\sqrt{\sigma_1^4 + \sigma_2^4}}{\sigma_1\sigma_2}$$

which is to some extent related to the coherence of the structure tensor (see section 3.4).

Finally, all fundamentals about the 2D Fourier transform which are needed for the derivation of a sophisticated 2D analytic signal are given. The classical approaches for a 2D analytic signal are to some extent unsatisfactory (see section 1.2 in the introduction). Therefore, I do not discuss their formal definitions in this section.[11]

Additionally to the presented fundamentals, I refer to further well known methods of 2D signal processing in chapter 5: the structure tensor and steerable filters. Both are introduced in the subsequent two sections.

3.4 The Structure Tensor

The new approaches which are presented in this thesis are related to both, the analytic signal and the structure tensor. While the coding of information by means

[11] Bracewell even omits them totally in [12], apparently since he is not pleased with their properties.

of the phase is taken from the analytic signal, the *geometric analysis* is taken from the structure tensor. Therefore, the latter is introduced in this section for the case of 2D signals.

3.4.1 The Structure Tensor and the Intrinsic Dimension

The structure tensor has been developed independently in at least two groups. The first references for tensor approaches are the paper of Förstner and Gülch [55] and the paper of Bigün and Granlund [7]. Nowadays, the structure tensor is one of the well known methods in multidimensional signal processing and it is explained in detail in [74, 61]. Since the structure tensor belongs to the methods making use of matrix algebra, I explain it in the classical framework, i.e., not in the vector field embedding. Hence, a 2D signal f is real valued in this section. Basically, there are two ways to compute the 2×2 matrix representing the tensor for 2D signals.

The first method is to calculate the gradient of the image $\nabla_2 f = (f_x, f_y)^{\mathrm{T}}$ and then to form the dyadic product:

$$\mathbf{T} = (\nabla_2 f)(\nabla_2 f)^{\mathrm{T}} = \begin{pmatrix} f_x f_x & f_x f_y \\ f_y f_x & f_y f_y \end{pmatrix} \quad . \tag{3.48}$$

According to [54], this matrix can be used for approximating the auto-covariance function of the signal in the origin. The structure tensor build from the gradient can also be related to the *tensor of inertia* [74].

The second method is to calculate four quadrature filter responses[12] in the directions $0°$, $45°$, $90°$, and 135, denoted by q_1, q_2, q_3, and q_4, respectively. In a second step, the amplitudes of the quadrature filters are combined as follows (see [79]):

$$\mathbf{T} = \frac{1}{4} \begin{pmatrix} 3|q_1| + |q_2| - |q_3| + |q_4| & 2|q_2| - 2|q_4| \\ 2|q_2| - 2|q_4| & -|q_1| + |q_2| + 3|q_3| + |q_4| \end{pmatrix} \quad . \tag{3.49}$$

The quadrature filters that form the structure tensor are built by a lognormal radial bandpass filter $B(|\mathbf{u}|)$ and the squared scalar product of frequency vector and directional vector in the positive half space of the directional vector [61]:

$$Q_k(\mathbf{u}) = \begin{cases} B(|\mathbf{u}|)(\mathbf{u} \cdot \mathbf{n}_k)^2 & \text{if } \mathbf{u} \cdot \mathbf{n}_k > 0 \\ 0 & \text{else} \end{cases} \quad .$$

The resulting angular function is a $\cos^2$-function which is centered at the angle of the directional vector (see figure 3.5). The key idea is to analyze the angular behavior of the local spectrum.

[12] Actually, the number of quadrature filters is not fixed to four. Any number of filters greater or equal to three is appropriate [61]. The construction of the tensor is especially simple in the case of four filters.

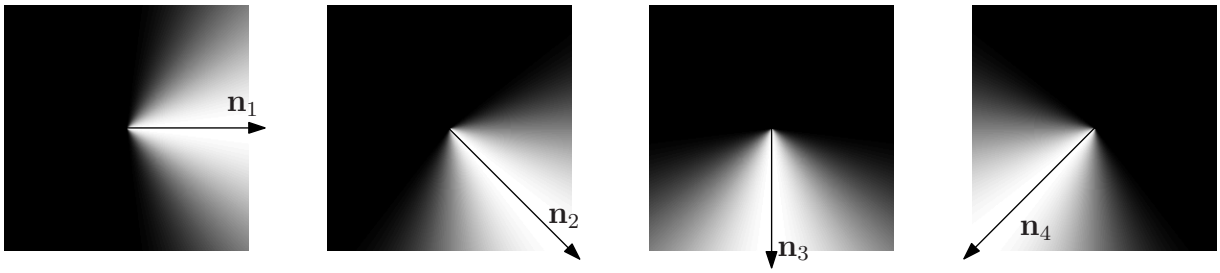

Fig. 3.5: Angular frequency responses of the quadrature filters used for the computation of the structure tensor according to (3.49).

For the following discussion, the notion of *intrinsic dimension* is needed [88, 9].

- Intrinsically zero dimensional (i0D) signals are constant in the spatial domain. Hence, there is no structure and their frequency representation is a delta function in the origin which indicates the DC-component (see left column of figure 3.6).
- Intrinsically one dimensional (i1D) signals are constant in one orientation in the spatial domain. Hence, their frequency representation is concentrated on a line which is orthogonal to the constant lines in the spatial domain (see middle column of figure 3.6, $\mathbf{n}$ indicates the vector which is normal to the constant lines).
- Intrinsically two dimensional (i2D) signals vary in two different orientations, and therefore in any orientation. Their frequency representation is not concentrated on a line (see right column of figure 3.6).

Obviously, the cardinality of the set of i2D signals is by orders higher than the cardinality of the set of i1D signals which is again much higher than the cardinality of the set of i0D signals.

In real images all three cases occur in different parts of the image. Hence, it is reasonable to define the *local intrinsic dimension*, where 'local' depends on the context, e.g. the scale of the used filters. Furthermore, any image can be decomposed into one i0D part (the DC-component) and i1D components – this is what the 2D Fourier transform does since its basis functions are all i1D functions, except for the frequency zero. The 2D Fourier transform is a *plane wave decomposition* of the image. This fact is fundamental to the analysis of i2D points, Radon transform theorems [119], and the new i1D approach (see chapter 4).

Both formulations of the structure tensor yield symmetric tensors which have two real eigenvalues in general. Except for constant signals, the eigenvalue decom-

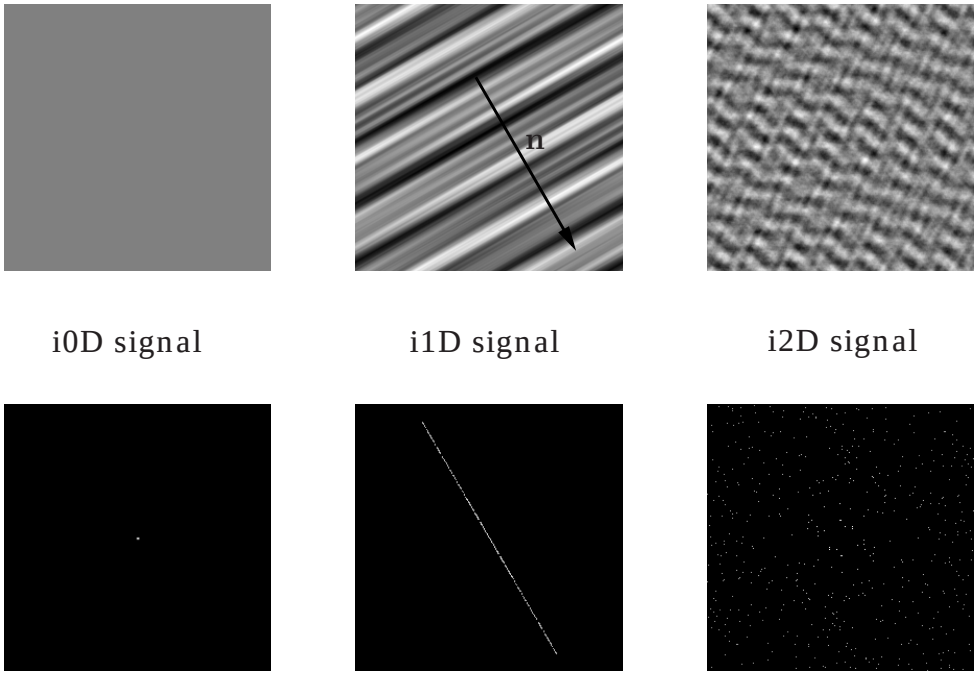

Fig. 3.6: From left to right: i0D signal, i1D signal, i2D signal. Upper row: spatial domain, bottom row: frequency domain. The vector $\mathbf{n}$ indicates the normal to the lines of constant value.

position of the structure tensor results at least in one non-zero eigenvalue[13] λ_1^2 with eigenvector $\mathbf{v}_1 = (\cos\theta_0, \sin\theta_0)^{\mathrm{T}}$. This first eigenvector represents the *local (main) orientation* θ_0 of the structure. The main orientation is only unique in an interval $[0, \pi)$ since if $\mathbf{v}_1$ is a unit eigenvector of $\mathbf{T}$, $-\mathbf{v}_1$ is also a unit eigenvector of $\mathbf{T}$ (see also [74]). Hence, it is more convenient to use the *double angle representation* [5, 59], i.e., the main orientation is represented by the vector $\mathbf{o} = |\mathbf{o}|(\cos(2\theta_0), \sin(2\theta_0))^{\mathrm{T}}$. This vector can be obtained directly from the components of the tensor $\mathbf{T}$ [8]:

$$\mathbf{o} = \begin{pmatrix} T_{11} - T_{22} \\ 2T_{12} \end{pmatrix} = \begin{cases} \begin{pmatrix} f_x^2 - f_y^2 \\ 2 f_x f_y \end{pmatrix} & \text{if } \mathbf{T} \text{ is obtained from (3.48)} \\ \begin{pmatrix} |q_1| - |q_3| \\ |q_2| - |q_4| \end{pmatrix} & \text{if } \mathbf{T} \text{ is obtained from (3.49)} \end{cases} \tag{3.50}$$

where $|\mathbf{o}| = \lambda_1^2 - \lambda_2^2$ is the difference of the eigenvalues. This result is obtained by plugging in $\lambda_1^2 > \lambda_2^2$, $\det(\mathbf{T}) = \lambda_1^2 \lambda_2^2$, and $\mathrm{trace}(\mathbf{T}) = \lambda_1^2 + \lambda_2^2$ (see e.g. [24] and footnote 13).

[13] In contrast to the common notation for eigenvalues, I use λ_k^2 instead of λ_k in order to be consistent with the dilation (3.36) and chapter 5. Furthermore, the eigenvalues of the structure tensor are non-negative (at least if the tensor is built according to (3.48)) and are a second order property.

Since it is not trivial that the vector $\mathbf{o}$ represents the main orientation, I will show this in the subsequent two paragraphs. If the signal is locally i1D, the local orientation is given by the normal vector $\mathbf{n}$ (see figure 3.6). Hence, the 2D signal is obtained by $f(\mathbf{x}) = g(\mathbf{x} \cdot \mathbf{n})$ where g is a 1D real valued function. Accordingly, $\mathbf{n} = (\cos\theta_0, \sin\theta_0)^{\mathrm{T}}$ and the spectrum of the signal is a Dirac line with angle θ_0 to the $\mathbf{e}_1$-axis.

For the structure tensor built according to (3.49), the quadrature filter responses are proportional to $\cos^2(\theta_0)$, $\cos^2(\theta_0 - \pi/4)$, $\cos^2(\theta_0 - \pi/2)$, and $\cos^2(\theta_0 - 3\pi/4)$. Combining these envelopes according to (3.50) yields

$$\mathbf{o} = |\mathbf{o}|(\cos^2(\theta_0) - \sin^2(\theta_0), 2\cos(\theta_0)\sin(\theta_0))^{\mathrm{T}} = |\mathbf{o}|(\cos(2\theta_0), \sin(2\theta_0))^{\mathrm{T}} \quad . \tag{3.51}$$

If the structure tensor is formed by the gradient of the image, the derivative theorem of the Fourier transform implies that the local spectrum is weighted by u and v. Due to the Plancherel theorem, averaging the products of the partial derivatives locally corresponds to the energy of the weighted local spectra:

$$\begin{aligned}\mathbf{T} &\sim \iint \begin{pmatrix} uF(\mathbf{u})uF^*(\mathbf{u}) & uF(\mathbf{u})vF^*(\mathbf{u}) \\ vF(\mathbf{u})uF^*(\mathbf{u}) & vF(\mathbf{u})vF^*(\mathbf{u}) \end{pmatrix} du\, dv \\ &= \iint \begin{pmatrix} u^2|F(\mathbf{u})|^2 & uv|F(\mathbf{u})|^2 \\ uv|F(\mathbf{u})|^2 & v^2|F(\mathbf{u})|^2 \end{pmatrix} du\, dv \quad .\end{aligned}$$

Plugging in $\mathbf{u} = q(\cos\theta, \sin\theta)^{\mathrm{T}}$ and taking into account that F is a Dirac line with orientation θ_0, (3.50) also yields the double angle vector $\mathbf{o} = |\mathbf{o}|(\cos(2\theta_0), \sin(2\theta_0))^{\mathrm{T}}$.

3.4.2 Interpretation of the Structure Tensor

Returning to the eigenvalue decomposition of the structure tensor, it is obvious that the tensor built according to (3.48) is of rank one, and hence its second eigenvalue is always zero. If the structure tensor is *averaged* over a local region however, the tensor is not restricted to be of rank one any more. The structure tensor according to (3.49) is of rank two in general. For tensors of rank two it is reasonable to consider both eigenvalues instead of looking only at their difference $|\mathbf{o}| = \lambda_1^2 - \lambda_2^2$. Since the trace of the tensor yields the sum of eigenvalues, both eigenvalues are obtained without a real decomposition of the tensor. Considering both eigenvalues, several cases can occur:

- The region only consists of noise. The coefficients of the tensor will probably cancel out and both eigenvalues are close to zero. The local region is i0D.
- The region is i1D. All orientation vectors point in the same direction and the averaging of the gradient tensor just yields a multiplication of all tensor coefficients by a constant, and hence the tensor according to (3.48) is of rank one.

The tensor according to (3.49) is of rank one either (see calculation preceding (3.51)). Under the presence of noise, the first eigenvalue is large and the second one is close to zero.

- The region is i2D. The tensor is of rank two. Therefore, both eigenvalues are obtained to be non-zero. The orientation vector is still given by (3.50), yielding the double angle representation of the first eigenvector $\mathbf{v}_1$. The second eigenvector is orthogonal to $\mathbf{v}_1$ where the directional sense is again indefinite.

Hence, if the second eigenvalue of the tensor is non-zero, this indicates an i2D neighborhood. If the structure tensor is interpreted in terms of an approximation of the auto-covariance function, the two eigenvalues correspond to the two principal curvatures (see e.g. [62] for the definition of principal curvature) of that function and the eigenvectors indicate the corresponding axes. On the other hand, the eigenvector of the larger eigenvalue indicates the orientation in which the variance of the derivative is maximal [54].

According to the construction of the tensor either by $\cos^2$-quadrature filters or by products of partial derivatives, the eigenvalues of the structure tensor can be considered as a $\cos^2$-*decomposition of the local amplitude* (see calculations in the subsequent paragraph). This corresponds basically to a Karhunen-Loéve transform [9, 113] of the local amplitude spectrum. If the structure tensor is formed by partial derivatives, the Karhunen-Loéve transform is applied to $q^2|F(\mathbf{u})|^2$. If the tensor is constructed from quadrature filters, the Karhunen-Loéve transform is applied to $B(q)|F(\mathbf{u})|$, where $B(q)$ is the frequency response of the radial bandpass.[14]

The $\cos^2$-decomposition is obtained as follows: Let θ_0 be the main orientation of the structure. The first eigenvector is then given by $\mathbf{v}_1 = (\cos\theta_0, \sin\theta_0)^{\mathrm{T}}$ and accordingly, the second one by $\mathbf{v}_2 = (-\sin\theta_0, \cos\theta_0)^{\mathrm{T}}$. According to the definitions of eigenvectors, it follows that

$$\mathbf{T}\mathbf{v}_k = \lambda_k^2 \mathbf{v}_k \quad \text{for } k = 1, 2 \ . \tag{3.52}$$

Multiplication with $\mathbf{v}_k^{\mathrm{T}}$ from the left gives

$$\mathbf{v}_k^{\mathrm{T}} \mathbf{T} \mathbf{v}_k = \lambda_k^2 \quad \text{for } k = 1, 2 \ .$$

Now let $\mathbf{T}$ be the (unit) tensor of a structure rotated by a small angle $\Delta\theta$. Hence, the two-sided multiplication by the eigenvector $\mathbf{v}_1$ gives

$$\begin{aligned} &\mathbf{v}_1^{\mathrm{T}} \begin{pmatrix} \cos^2(\theta_0 + \Delta\theta) & \cos(\theta_0 + \Delta\theta)\sin(\theta_0 + \Delta\theta) \\ \cos(\theta_0 + \Delta\theta)\sin(\theta_0 + \Delta\theta) & \sin^2(\theta_0 + \Delta\theta) \end{pmatrix} \mathbf{v}_1 \\ &\qquad = \mathbf{v}_1^{\mathrm{T}} \begin{pmatrix} \cos(\theta_0 + \Delta\theta) \\ \sin(\theta_0 + \Delta\theta) \end{pmatrix} (\cos(\theta_0 + \Delta\theta)\,,\, \sin(\theta_0 + \Delta\theta))\, \mathbf{v}_1 = \cos^2 \Delta\theta \ . \end{aligned}$$

[14] This is only valid in its full extent if the signal is i1D. Otherwise the tensor is an approximation of the Karhunen-Loéve transform.

And similarly

$$\mathbf{v}_2^{\mathrm{T}} \begin{pmatrix} \cos(\theta_0 + \Delta\theta) \\ \sin(\theta_0 + \Delta\theta) \end{pmatrix} (\cos(\theta_0 + \Delta\theta)\,,\, \sin(\theta_0 + \Delta\theta))\, \mathbf{v}_2 = \sin^2 \Delta\theta \quad .$$

Hence, having several local structures, each one contributes to the first eigenvalue weighted by the cosine-square of the angular distance to the main orientation. The contribution to the second eigenvalue is given by the corresponding sine-square.

Since $\cos^2 \Delta\theta = \frac{1}{2}(1 + \cos(2\Delta\theta))$ and $\sin^2 \Delta\theta = \frac{1}{2}(1 - \cos(2\Delta\theta))$, the sum of both eigenvalues is *independent of the angular distribution*, while their difference yields a measure for the angular variance. Hence, the *coherence* of the local structure is defined by [74] as

$$c = \left(\frac{\lambda_1^2 - \lambda_2^2}{\lambda_1^2 + \lambda_2^2} \right) = \left(\frac{|\mathbf{o}|}{\mathrm{trace}(\mathbf{T})} \right) \quad , \tag{3.53}$$

which is independent of the local dynamics of the signal. Assumed that the denominator is non-zero, the coherence is of value one if the underlying structure is i1D, and it decreases with raising intensity of i2D structures. For an 'ideal' i2D structure (the energy is distributed uniformly for all angles), the coherence is zero. Hence, the coherence can be used to design a *corner detector* or *point-of-interest operator*. Using the coherence in a slightly different formulation (the eigenvalues are weighted) yields the *Plessey corner detector* which is also named Stephens detector or Harris-Stephens detector [68]. Besides the two extreme cases of i1D and i2D structures, the coherence provides a mean for estimating the curvature of arc-like structure since it is a smooth, monotonic map of the angular variance. Consider e.g. two harmonic oscillations with equal amplitudes and with frequency vectors which differ only by their angle. The coherence is then basically given by the cosine of the angle difference.

The detection of i2D points is based on the energy distribution of the local structure. Hence, special kinds of structure are identified by means of energy. Opposed to that, the analytic signal of a 1D signal encodes the structure in its phase (see section 3.2). Moreover, the structure tensor does not include any information about the symmetry of i1D structures due to the lack of a phase concept. In chapter 4 an approach is presented which evaluates the local phase of an i1D structure. Basically, the phase is obtained from the local phase of the underlying 1D signal. Obviously, such a phase concept is insufficient to describe i2D structures. Therefore, the 2D local phase cannot be a single phase angle like the phase in 1D, but it must be split into *i1D 2D phase* and *i2D 2D phase*. The i1D 2D phase is directly given by the underlying 1D signal. The design of a i2D 2D phase is not obvious in the light of the current knowledge, it has to be developed based on geometric ideas.

3.5 Spherical Harmonics and Steerable Filters

In contrast to the Fourier transform, the *Fourier series* applies to functions on periodic or finite domains[15]. In [140] a mathematical classification of different cases of domains can be found which is based on group theory. In its most common formulation, the Fourier series is defined for arbitrary periodic functions. In the present section, the considerations are restricted to 2π-periodic functions.

3.5.1 Fourier Series

To be more specific, the 'spatial' coordinate is an angle θ which is obtained from the 2D spatial vector $\mathbf{x}$ according to

$$\theta \mathbf{e}_{12} = \langle \log(\mathbf{e}_1 \mathbf{x}) \rangle_2 \ , \tag{3.54}$$

so that $\mathbf{x} = |\mathbf{x}|(\cos\theta \mathbf{e}_1 + \sin\theta \mathbf{e}_2)$. Knowing that angles are bivectors (see also section 2.2), it is not necessary to introduce an additional imaginary unit, since it is directly given by $\mathbf{e}_{12}$. Therefore, the definition of the Fourier series of an angular, $\mathbb{R}_2^+$-valued[16] function h is given by

$$h(\theta) = \sum_{n \in \mathbb{Z}} H_n \exp(n\theta \mathbf{e}_{12}) \ , \tag{3.55}$$

or in terms of (3.54) by

$$h\left(\log\left(\frac{\mathbf{e}_1 \mathbf{x}}{|\mathbf{x}|}\right)\right) = \sum_{n \in \mathbb{Z}} H_n \left(\frac{\mathbf{e}_1 \mathbf{x}}{|\mathbf{x}|}\right)^n \ . \tag{3.56}$$

The basis functions $\exp(n\theta \mathbf{e}_{12})$ are called *spherical harmonics*[17] of order n and the coefficients H_n of the series are obtained by

$$H_n = \frac{1}{2\pi} \int_{\theta=0}^{2\pi} h(\theta) \exp(-n\theta \mathbf{e}_{12}) \, d\theta \ . \tag{3.57}$$

If the series (3.55) is truncated by $|n| < n_{\max}$, the series is still a good approximation of $h(\theta)$. Indeed, one can show that the Fourier series is the *optimal* approximation with respect to the L^2-norm (see [84], chapter 10).

[15] Formally, there is no difference between these two cases since finite domains can be viewed as periodic.

[16] The considered functions are restricted to $\mathbb{R}_2^+$-valued functions since the Fourier kernel commutes with these functions which makes the subsequent derivations much easier. The $\mathbb{R}_2^+$-valued functions are sufficient for the subsequent calculations in this thesis.

[17] The term spherical harmonic is more general. It denotes the harmonic oscillations on the unit sphere of a multi-dimensional Euclidean space. In this thesis however, solely the spherical harmonics of 2D space are used.

In the light of the L^2-optimality, consider the following situation. The orientation of a 2D signal has to be estimated. If the 2D signal is not i1D, the orientation is not given by a unique orientation vector. Suppose further that the radial bandwidth of the signal is small so that its 2D Fourier representation can be approximated solely by an angular function. This angular function is optimally approximated by the first order spherical Fourier series. Hence, the L^2-optimal estimate of the orientation is obtained from the first order spherical harmonic (see also chapter 4).

Returning to the Fourier series, it is obvious that there exists a counterpart to the shift-theorem of the Fourier transform since

$$\begin{aligned}\frac{1}{2\pi}\int_{\theta=0}^{2\pi} h(\theta-\theta_0)\exp(-n\theta\mathbf{e}_{12})\,d\theta \\ &= \frac{1}{2\pi}\int_{\theta=0}^{2\pi} h(\theta')\exp(-n\theta'\mathbf{e}_{12})\,d\theta'\,\exp(-n\theta_0\mathbf{e}_{12}) \\ &= \exp(-n\theta_0\mathbf{e}_{12})H_n\ . \end{aligned} \tag{3.58}$$

Hence, a rotation of h by θ_0 yields a modulation of the Fourier coefficients by the factor $\exp(-n\theta_0\mathbf{e}_{12})$. This effect can also be used to rotate signals or filters. Multiplying the Fourier coefficients of the original signal or filter with appropriate exponential factors yields the Fourier coefficients of a rotated version of the signal or filter. In the context of filter design, this method is known as *steerable filters*.

3.5.2 Applying Steerable Filters

The idea of steerable filters is the following: Since 2D convolutions are computational expensive (at least for large kernels), it is desirable to have a small set of basis filters which can be used to create more complex filters. If a filter has to be applied for various angles, it can either be applied for a large set of angles (sampling of the orientation) which yields a very high computational load, or it can be decomposed into its spherical Fourier series and each spherical harmonic contained in the series is convolved with the image. The filter response for the angle θ_0 can be calculated from the convolutions by the following sum:

$$\begin{aligned} g_{\theta_0}(\mathbf{x}) &= (h_{\theta_0} * f)(\mathbf{x}) \\ &= \left(\left(\sum_{n\in\mathbb{Z}} \exp(-n\theta_0\mathbf{e}_{12})H_n\exp(n\theta\mathbf{e}_{12})\right) * f\right)(\mathbf{x}) \\ &= \sum_{n\in\mathbb{Z}} \exp(-n\theta_0\mathbf{e}_{12})H_n(\exp(n\theta\mathbf{e}_{12}) * f)(\mathbf{x}) \end{aligned} \tag{3.59}$$

(see also [57] and for a more general approach [101, 102]). Obviously the steerable filter approach only yields a computational advantage if the Fourier series of the

filter is finite ($H_n = 0$ for $|n| > n_{\max}$) and $n_{\max}$ is small. In the case of the partial Hilbert transform (see section 1.2), the Fourier series is infinite, since it is basically the series of an angular rectangle in the Fourier domain [16]:

$$H_n = \begin{cases} \frac{2}{\pi n} & n \text{ is odd} \\ 0 & n \text{ is even} \end{cases} . \tag{3.60}$$

Hence, the Hilbert transform can only be *approximated* by a steerable filter. This is also true for other useful filters, see e.g. [132, 2]. The problem in the context of steering the Hilbert transform is however, that the approximation error introduces further phase distortions additional to those which are always present in the case of the partial Hilbert transform.

Using steerable filters is also advantageous in the case of *orientation adaptive* filtering, where the orientation used for steering the filter response is evaluated from the signal itself. Orientation adaptive filtering yields a method for processing the signal efficiently in a *semi-linear* and *shift-variant* way. The method is efficient because the basis filters are LSI filters and the shift-variance is not introduced before the subsequent steering operation. Semi-linearity means, that the approach is linear for signals with the same orientation, whereas this is not certain in general.

Herewith I conclude this chapter about signal theoretic fundamentals. In the subsequent chapter, a new method for generalizing the Hilbert transform is introduced which does not show any orientation dependent phase distortions and which is linear and shift-invariant, i.e., no steering is applied. For the detection of i2D structures however, steering is necessary (see chapter 5).

3.6 Summary of Chapter 3

In order to prepare for the subsequent chapters, 1D and 2D signal theory has been rewritten in terms of geometric algebra in this chapter. In contrast to the commonly used formulation in complex numbers, geometric algebra allows to distinguish between objects as vectors and operations as spinors. Accordingly, signals are embedded as vector valued functions whereas filters are embedded as spinor valued functions. This embedding yields an interpretation of nD signals as vector fields in $(n+1)$D. The formulation of the Fourier transform and the related theorems is slightly changed, although the fundamental relationships remain the same. However, the decisive advantage of the embedding is that the 1D analytic signal can be interpreted as a harmonic field which will be used in the subsequent chapter for defining its 2D generalization. The embedding of 1D and 2D signal theory is summarized in table 3.1 and table 3.2.

Besides the reformulation of signal theory, the term of the intrinsic dimension has been introduced which is given by the dimension of the (sub-) space in which

dimension	1	2
Fourier transf.	$G(\mathbf{u}) = \int \exp(-\mathbf{e}_{12} 2\pi \mathbf{x}\mathbf{u}) g(\mathbf{x})\, dx$	$G(\mathbf{u}) = \iint g(\mathbf{x}) \exp(-I_3 2\pi \mathbf{x} \cdot \mathbf{u})\, d\mathbf{x}$
shift theorem	$\mathcal{F}_1\{\delta_0(\mathbf{x}-\mathbf{x}')\} = \exp(-\mathbf{e}_{12} 2\pi \mathbf{x}'\mathbf{u})$	$\mathcal{F}_2\{\delta_0(\mathbf{x}-\mathbf{x}')\} = \exp(-I_3 2\pi \mathbf{x}' \cdot \mathbf{u})$
gradient	$\mathcal{F}_1\{\nabla_1 g(\mathbf{x})\} = 2\pi \mathbf{u}^* G(-\mathbf{u})$	$\mathcal{F}_2\{\nabla_2 g(\mathbf{x})\}(\mathbf{u}) = -2\pi \mathbf{u}^* G(\mathbf{u})$
Laplace	$\mathcal{F}_1\{\Delta_1\}(\mathbf{u}) = -4\pi^2 \mathbf{u}^2$	$\mathcal{F}_2\{\Delta_2\}(\mathbf{u}) = -4\pi^2 \mathbf{u}^2$
uncertainty	$(\Delta \mathbf{x})(\Delta \mathbf{u}) \geq \frac{1}{4\pi}$	$(\Delta \mathbf{x})(\Delta \mathbf{u}) \geq \frac{1}{2\pi}$
phase	$\arg(H(\mathbf{u})) = \langle \log(H(\mathbf{u})) \rangle_2$	see chapter 4
amplitude	$\lvert H(\mathbf{u}) \rvert = \exp(\langle \log(H(\mathbf{u})) \rangle_0)$	see chapter 4
analytic signal	$\mathbf{f}_\mathrm{A}(\mathbf{x}) = \mathbf{f}(\mathbf{x}) + \mathbf{f}_\mathrm{H}(\mathbf{x})$	see chapter 4
Hilbert transf.	$h_1(\mathbf{x}) = \frac{\mathbf{x}}{\pi \lvert \mathbf{x} \rvert^2} \mathbf{e}_2$ / $H_1(\mathbf{u}) = \mathbf{e}_1 \frac{\mathbf{u}}{\lvert \mathbf{u} \rvert}$	see chapter 4

Tab. 3.1: Embedding of 1D and 2D signal theory in geometric algebra.

name	formulation	remarks
affine 1	$\mathcal{F}_2\{g(\tilde{s}\mathbf{x}s)\}(\mathbf{u}) = \lambda^{-2} G(s^{-1}\mathbf{u}\tilde{s}^{-1})$	$s = \sqrt{\lambda} \exp(\mathbf{e}_{12}\theta/2)$
affine 2	$\mathcal{F}_2\{g(\mathbf{x}')\} = \frac{1}{\lambda_1 \lambda_2} G(\mathbf{u}')$	$\mathbf{x}' = \lambda_1 \mathbf{n}(\mathbf{n} \cdot \mathbf{x}) + \lambda_2 \mathbf{n}(\mathbf{n} \wedge \mathbf{x})$ and $\mathbf{u}' = \lambda_1^{-1} \mathbf{n}(\mathbf{n} \cdot \mathbf{u}) + \lambda_2^{-1} \mathbf{n}(\mathbf{n} \wedge \mathbf{u})$
Plancherel	$\langle\langle g_1, g_2 \rangle\rangle_k = \langle\langle G_1, G_2 \rangle\rangle_k$	implies Parseval theorem
uncertainty	$((\Delta \mathbf{x})(\Delta \mathbf{u}))^{\parallel} + ((\Delta \mathbf{x})(\Delta \mathbf{u}))^{\perp} \mathbf{e}_{12}$	signal is separable

Tab. 3.2: Theorems of 2D signal theory in geometric algebra.

the signal varies. The structure tensor is the appropriate tool for estimating the intrinsic dimension if the focus lies on the signal amplitudes solely. By means of an eigenvalue decomposition of the tensor, the local orientation, the local amplitude, and the local coherence of a signal are estimated. These three signal features fulfill the invariance – equivariance requirement as it is the case for quadrature filters. However, no local phase approach is contained in the structure tensor.

As a last topic of this chapter, the method of steerable filters has been briefly explained. It is based on a truncated series of spherical harmonics and the steering of the orientation is obtained by means of the shift theorem. However, signals which have no angular band-limitation can only be approximated by a steerable filter, as it is the case for the partial Hilbert transform.

Chapter 4

THE ISOTROPIC 2D GENERALIZATION OF THE ANALYTIC SIGNAL

'The further a mathematical theory is developed,
the more harmoniously and uniformly
does its construction proceed,
and unsuspected relations are disclosed between
hitherto separated branches of the science.'

David Hilbert (1862-1943)

At the end of section 3.4, I suggested to design different 2D phase approaches for different intrinsic dimensions. For i1D structures, one should keep the phase concept from the 1D analytic signal. Hence, derivations of a new 2D analytic signal should be related to the 1D case. In this section I therefore present a method to derive the Hilbert transform and the analytic signal by means of the vector field embedding described in section 3.1.

4.1 Derivation of the Hilbert Transform

The Hilbert transform of 1D signals emerges from complex analysis by means of the Cauchy formula [66]. The basic idea of the derivation is that the contour integral of the Cauchy formula is replaced by an integral on the real line which is basically the Hilbert transform. Since complex analysis and in particular its multidimensional generalizations are not very illustrative and therefore sometimes hard to understand, I make use of vector field theory instead. It is a well known fact that *analytic functions* of one complex variable directly correspond to 2D harmonic fields [86] (see also section 2.3). This generalizes to 3D such that *monogenic functions* correspond to 3D harmonic fields (see section 2.4). This equivalence is used in section 4.2 to de-

rive the Riesz transform as a 2D generalization of the Hilbert transform.[1] The aim of this section is to derive the Hilbert transform by means of 2D harmonic fields. I will show that the Hilbert transform relates the components of such a vector field on every line with arbitrary, but fixed $y = y_0$. The main results from this section are summarized in figure 4.2.

4.1.1 The 1D Signal Model

According to chapter 3, a real 1D signal is embedded as an $\mathbf{e}_2$-valued vector field on the line $y = 0$. Establishing then a boundary value problem of the second kind (see (2.49) and (2.50)) is equivalent to finding the holomorphic extension of the real signal. The holomorphic extension restricted to the line $y = 0$ is just the analytic signal (see also figure 4.1).

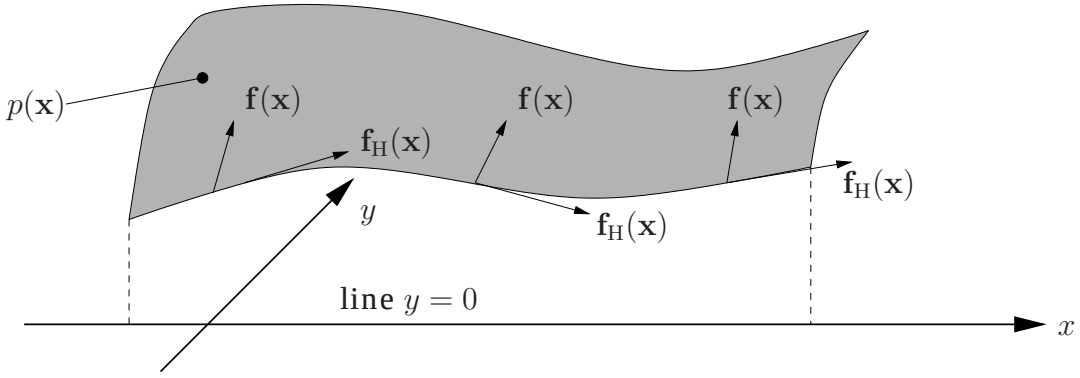

Fig. 4.1: Harmonic potential $p(\mathbf{x})$ for $y > 0$ and its gradient field on the boundary $y = 0$, which is given by the 1D function $\mathbf{f}(\mathbf{x})$ and its Hilbert transform $\mathbf{f}_{\mathrm{H}}(\mathbf{x})$.

In contrast to the derivation in [46], I do not focus on the relationship between $\mathbf{f}$ and $\mathbf{f}_{\mathrm{H}}$ exclusively. Instead, the fundamental solution (2.48) of the Laplace equation is used. The $\mathbf{e}_2$-component of the harmonic field is given for $y = 0$ by $\mathbf{f}$. The $\mathbf{e}_2$-component for $y = y_0 > 0$ is obtained from the y-derivative of $p_2\mathbf{e}_2$,[2]

$$(\mathbf{f}_2 \cdot \mathbf{e}_2)(\mathbf{x}) = \mathbf{e}_2 \frac{\partial}{\partial y} p_2(\mathbf{x})\mathbf{e}_2 = \frac{y}{\pi|\mathbf{x}|^2} \quad , \tag{4.1}$$

[1] The relationship between harmonic functions of $\mathbb{R}^n \oplus \mathbb{R}^+$ and singular operators in $\mathbb{R}^n$ is quite fundamental, see [138] chapter III.4

[2] In this context the potential p_2 is used to obtain the scalar valued potential with the y-derivative $\mathbf{f}$ for $y = 0$. Since $\mathbf{f}$ is $\mathbf{e}_2$-valued, p_2 must be multiplied by $\mathbf{e}_2$ to obtain a scalar function. In mathematical terms $\mathbf{e}_2$ results from the measure $d\sigma(x) = \mathbf{e}_2\, dx$ (see [58], page 121).

which is known as the *Poisson kernel* [138]. Embedding the 1D signal f into $\mathbb{R}^2$ according to $\mathbf{f}_0(\mathbf{x}) = \mathbf{f}(x\mathbf{e}_1)\delta_0(y)$, the following 2D convolution extends f to the whole half-space $y > 0$:

$$\begin{aligned}\mathbf{f}(\mathbf{x}) = ((\mathbf{f}_2 \cdot \mathbf{e}_2) * \mathbf{f}_0)(\mathbf{x}) &= \iint \frac{y'}{\pi|\mathbf{x}'|^2}\mathbf{f}((x-x')\mathbf{e}_1)\delta_0(y-y')\,dx'\,dy' \\ &= \int \frac{y}{\pi|x'\mathbf{e}_1 + y\mathbf{e}_2|^2}\mathbf{f}((x-x')\mathbf{e}_1)\,dx' \quad . \end{aligned} \tag{4.2}$$

Since the 2D convolution is reduced to a 1D convolution, the latter can also be evaluated in the frequency domain of x. The 1D Fourier transform of the Poisson kernel is given in e.g. [136], chapter I (for the sake of clarity I use the components of spatial and frequency vectors instead of the vectors themselves):

$$\mathcal{F}\{\mathbf{f}_2 \cdot \mathbf{e}_2\}(u,y) = \exp(-2\pi|u|y) \tag{4.3}$$

(see also appendix A.1). Hence,

$$\mathcal{F}\{\mathbf{f}\}(u,y) = \exp(-2\pi|u|y)\mathbf{F}(u\mathbf{e}_1) \tag{4.4}$$

is the 1D Fourier transform of the extension of f to the whole half-space $y > 0$. Obviously, the Poisson kernel is a *lowpass filter* where the degree of smoothness increases with y. Therefore, y plays the role of a *scale parameter* and by the Poisson kernel a linear scale-space is established. This is discussed more in detail in section 4.5.

According to figure 4.1 it has to be shown that the $\mathbf{e}_1$-component of the considered field is the Hilbert transform of the corresponding $\mathbf{e}_2$-component for any fixed $y_0 \geq 0$. The x-derivative of $p_2\mathbf{e}_2$ is given by

$$(\mathbf{f}_2 \wedge \mathbf{e}_2)(\mathbf{x}) = \mathbf{e}_1\frac{\partial}{\partial x}p_2(\mathbf{x})\mathbf{e}_2 = \frac{x}{\pi|\mathbf{x}|^2}I_2 \quad , \tag{4.5}$$

which is called the *conjugate Poisson kernel* [138]. Applying the Fourier transform yields

$$\mathcal{F}\{\mathbf{f}_2 \wedge \mathbf{e}_2\}(u,y) = \frac{u}{|u|}\exp(-2\pi|u|y) \quad , \tag{4.6}$$

(see also appendix A.1) so that

$$\frac{u}{|u|}\exp(-2\pi|u|y)\mathbf{F}(u\mathbf{e}_1) = \exp(-2\pi|u|y)\mathbf{F}_{\mathrm{H}}(u) \quad , \tag{4.7}$$

which is the extension of the Hilbert transform $\mathbf{f}_{\mathrm{H}}$ to the whole half-space $y > 0$ (compare to equations (3.22) and (3.23)). The relationships just described are illustrated in figure 4.2.

The combination of the (smoothed) signal with its Hilbert transform (i.e., the vector field on a line $y = y_0$) forms the analytic signal for a certain scale y_0. For $y_0 = 0$ this is just the analytic signal $\mathbf{f}(\mathbf{x}) + \mathbf{f}_{\mathrm{H}}(\mathbf{x})$ (see (3.17)).

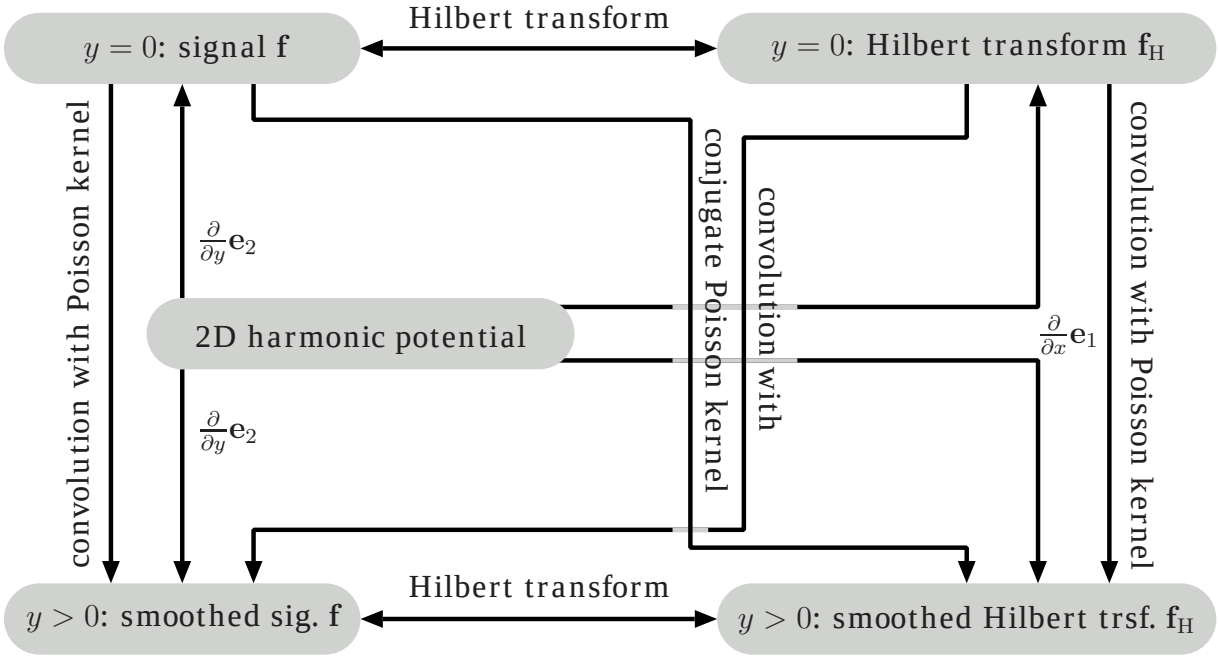

Fig. 4.2: Relationships between harmonic potential, signal, Hilbert transform, and convolution with the Poisson kernel.

4.1.2 Relation to the Laplace Transform

The combination of the Poisson kernel with its conjugate is closely related to the Laplace transform (see e.g. [127]), which is defined by

$$\mathcal{L}\{g(x)\}(p) = \int_0^\infty \exp(-px)g(x)\,dx \tag{4.8}$$

with $p \in \mathbb{R}_2^+ \simeq \mathbb{C}$ and $\langle p \rangle_0 > 0$. The spectrum of an analytic signal is one-sided in the frequency domain. Exchanging the roles of spatial (or time) domain and frequency domain results in a *causal* complex signal. Multiplying with the frequency response of the Poisson kernel and subsequent (inverse) Fourier transform is identical to the Laplace transform. Hence, *the holomorphic extension of a real valued signal in the spatial domain is equivalent to the Laplace transform of its Fourier transform*:

$$\begin{aligned}
\int_{-\infty}^{\infty} \frac{y + x' I_2}{\pi(x'^2 + y^2)} g(x - x')\,dx &= \int_{-\infty}^{\infty} \exp(I_2 2\pi u x)(1 + \operatorname{sign}(u)) \exp(-2\pi|u|y) G(u)\,du \\
&= 2\int_0^{\infty} \exp(I_2 2\pi u x) \exp(-2\pi u y) G(u)\,du \\
&= 2\int_0^{\infty} \exp(-(y - xI_2)2\pi u) G(u)\,du \\
&= 2\mathcal{L}\{G(u)\}(2\pi(y - xI_2)) \quad .
\end{aligned}$$

Therefore, the Laplace transform of the signal in frequency space yields its lowpass filtered analytic signal. The limit $y \to 0$ does always exist since the (inverse) Fourier transform of the spectrum exists. By means of this knowledge, the analytic signal can be analyzed solely by its poles and zeros in the xy-plane. However, a further discussion of this topic is out of the scope of this thesis.

4.2 The Riesz Transform

In the light of the previous section, I apply exactly the same steps in order to derive the 2D generalization of the Hilbert transform. The main results of this section are summarized in figure 4.4 which is basically the same as figure 4.2, but with the important difference that the Hilbert transform is replaced by the Riesz transform. The derived frequency responses and impulse responses are summarized in table 4.1.

4.2.1 The 2D Signal Model

According to chapter 3, a real 2D signal is embedded as an $\mathbf{e}_3$-valued vector field $\mathbf{f}$ in the plane $z = 0$. Establishing then a boundary value problem of the second kind (see (2.59) and (2.60)) is equivalent to finding the monogenic extension of the real signal.[3] The monogenic extension restricted to the plane $z = 0$ is then the 2D generalization of the analytic signal. The embedding of the boundary value problem is illustrated in figure 4.3.

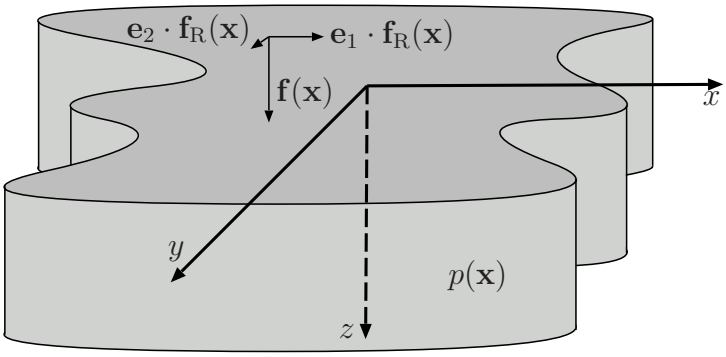

Fig. 4.3: Harmonic potential $p(\mathbf{x})$ for $z > 0$ and its gradient field on the boundary $z = 0$, which is given by the 2D function $\mathbf{f}_3(\mathbf{x})$ and its Riesz transform $\mathbf{f}_\mathrm{R}(\mathbf{x})$.

[3] Same as the holomorphic extension, the monogenic extension and the solution of the Neumann problem are unique [104], page 123.

As in the previous section, the fundamental solution (2.61) of the Laplace equation is used. The $\mathbf{e}_3$-component for $z = z_0 > 0$ is obtained from the z-derivative of $p_3\mathbf{e}_3$[4]

$$(\mathbf{f}_3 \cdot \mathbf{e}_3)(\mathbf{x}) = \mathbf{e}_3 \frac{\partial}{\partial z} p_3(\mathbf{x})\mathbf{e}_3 = \frac{z}{2\pi|\mathbf{x}|^3} \ , \tag{4.9}$$

which is known as the Poisson kernel for 2D functions. Embedding the 2D signal f into $\mathbb{R}^3$ according to $\mathbf{f}_0(\mathbf{x}) = \mathbf{f}(x\mathbf{e}_1 + y\mathbf{e}_2)\delta_0(z)$, the following 3D convolution extends f to the whole half-space $z > 0$:

$$\begin{aligned} \mathbf{f}(\mathbf{x}) &= ((\mathbf{f}_3 \cdot \mathbf{e}_3) * \mathbf{f}_0)(\mathbf{x}) \\ &= \iiint \frac{z'}{2\pi|\mathbf{x}'|^3} \mathbf{f}((x-x')\mathbf{e}_1 + (y-y')\mathbf{e}_2)\delta_0(z-z')\, dx'\, dy'\, dz' \\ &= \iint \frac{z}{2\pi|x'\mathbf{e}_1 + y'\mathbf{e}_2 + z\mathbf{e}_3|^3} \mathbf{f}((x-x')\mathbf{e}_1 + (y-y')\mathbf{e}_2)\, dx'\, dy' \ . \end{aligned} \tag{4.10}$$

Since the 3D convolution is reduced to a 2D convolution, the latter can also be evaluated in the frequency domain of x and y. The 2D Fourier transform of the Poisson kernel is given in e.g. [136] chapter I (for the sake of clarity I use again the components of spatial and frequency vectors instead of the vectors themselves):

$$\mathcal{F}\{\mathbf{f}_3 \cdot \mathbf{e}_3\}(u, v, z) = \exp(-2\pi|u\mathbf{e}_1 + v\mathbf{e}_2|z) \tag{4.11}$$

(see also appendix A.2). Hence,

$$\mathcal{F}\{\mathbf{f}\}(u, v, z) = \exp(-2\pi|u\mathbf{e}_1 + v\mathbf{e}_2|z)\mathbf{F}(u\mathbf{e}_1 + v\mathbf{e}_2) \tag{4.12}$$

is the 2D Fourier transform of the extension of f to the whole half-space $z > 0$. Obviously, the Poisson kernel is a *lowpass filter* where the degree of smoothness increases with z. Therefore, z plays the role of a *scale parameter* and by the Poisson kernel a linear scale-space is established. This is discussed more in detail in section 4.5.

According to figure 4.3 it has to be shown that the $\mathbf{e}_1$-component and the $\mathbf{e}_2$-component of the field above are obtained by a generalized Hilbert transform of the corresponding $\mathbf{e}_3$-component for a fixed $z_0 \geq 0$. The x-derivative and the y-derivative of $p_3\mathbf{e}_3$ are given by

$$(\mathbf{f}_3 \wedge \mathbf{e}_3)(\mathbf{x}) = \mathbf{e}_1 \frac{\partial}{\partial x} p_3(\mathbf{x})\mathbf{e}_3 + \mathbf{e}_2 \frac{\partial}{\partial y} p_3(\mathbf{x})\mathbf{e}_3 = -\frac{x}{2\pi|\mathbf{x}|^3}\mathbf{e}_{31} + \frac{y}{2\pi|\mathbf{x}|^3}\mathbf{e}_{23} \ , \tag{4.13}$$

which are the conjugate Poisson kernels [138]. Their Fourier transforms are (see appendix A.2)

$$\mathcal{F}\{\mathbf{f}_3 \wedge \mathbf{e}_3\}(u, v, z) = -\left(\frac{u\mathbf{e}_{31} - v\mathbf{e}_{23}}{|u\mathbf{e}_1 + v\mathbf{e}_2|}\right)^* \exp(-2\pi|u\mathbf{e}_1 + v\mathbf{e}_2|z) \ , \tag{4.14}$$

[4] Footnote 2 applies accordingly.

so that the frequency response

$$H_2(u\mathbf{e}_1 + v\mathbf{e}_2) = -\left(\frac{u\mathbf{e}_{31} - v\mathbf{e}_{23}}{|u\mathbf{e}_1 + v\mathbf{e}_2|}\right)^* = \frac{u\mathbf{e}_1 + v\mathbf{e}_2}{|u\mathbf{e}_1 + v\mathbf{e}_2|} I_2^{-1} \tag{4.15}$$

establishes the *Riesz transform* as a generalized Hilbert transform.[5]

At the end of this section several remarks about the Riesz transform, its origin, and its use in image processing are made. The Riesz transform is a singular integral operator as the Hilbert transform, but in contrast to the latter, it consists of two parts. Hence, for a 2D signal f, a vector consisting of two additional signals is obtained, which is denoted by $\mathbf{f}_\mathrm{R}$.

Corresponding to (4.7) in the 1D case,

$$\frac{u\mathbf{e}_1 + v\mathbf{e}_2}{|u\mathbf{e}_1 + v\mathbf{e}_2|} I_2^{-1} \exp(-2\pi|u\mathbf{e}_1 + v\mathbf{e}_2|z)\mathbf{F}(u\mathbf{e}_1 + v\mathbf{e}_2)$$
$$= \exp(-2\pi|u\mathbf{e}_1 + v\mathbf{e}_2|z)\mathbf{F}_\mathrm{R}(u\mathbf{e}_1 + v\mathbf{e}_2) \ , \tag{4.16}$$

is the extension of the Riesz transform $\mathbf{f}_\mathrm{R}$ to the whole half-space $z > 0$. The relationships just described are illustrated in figure 4.4.

The combination of the (smoothed) signal with its Riesz transform (i.e., the vector field on a line $z = z_0$) forms the generalized analytic signal for a certain scale. This generalization is analyzed more in detail in the subsequent section.

The properties and the geometric interpretation of the Riesz transform are also discussed in section 4.3. Note that from this section on, the z-coordinate is kept distinct from the spatial or the frequency coordinates. Hence, $\mathbf{x} = x\mathbf{e}_1 + y\mathbf{e}_2$, $\mathbf{u} = u\mathbf{e}_1 + v\mathbf{e}_2$, and z is replaced by s. The considered functions are no longer 3D functions, but 2D functions with parameter s, attached to the function symbol as a subscript (e.g. $\mathbf{f}_s$). Using the new notation, the previous results are summarized in table 4.1.

Before proceeding with the interpretation of the Riesz transform in the context of a generalized analytic signal, I want to refer to two aspects: the relation between the Fourier transform and the Laplace equation and the history of the Riesz transform.

4.2.2 Relation to the Fourier Kernel and a Short Survey

Instead of using the fundamental solution of the Laplace equation, the latter can also be solved directly by separation. The first step is to separate the s-derivative from the spatial derivatives:

$$\Delta_2 p_s(\mathbf{x}) = -\frac{\partial^2}{\partial s^2} p_s(\mathbf{x}) \ . \tag{4.23}$$

[5] In some places this operator is also called the Cauchy operator [94].

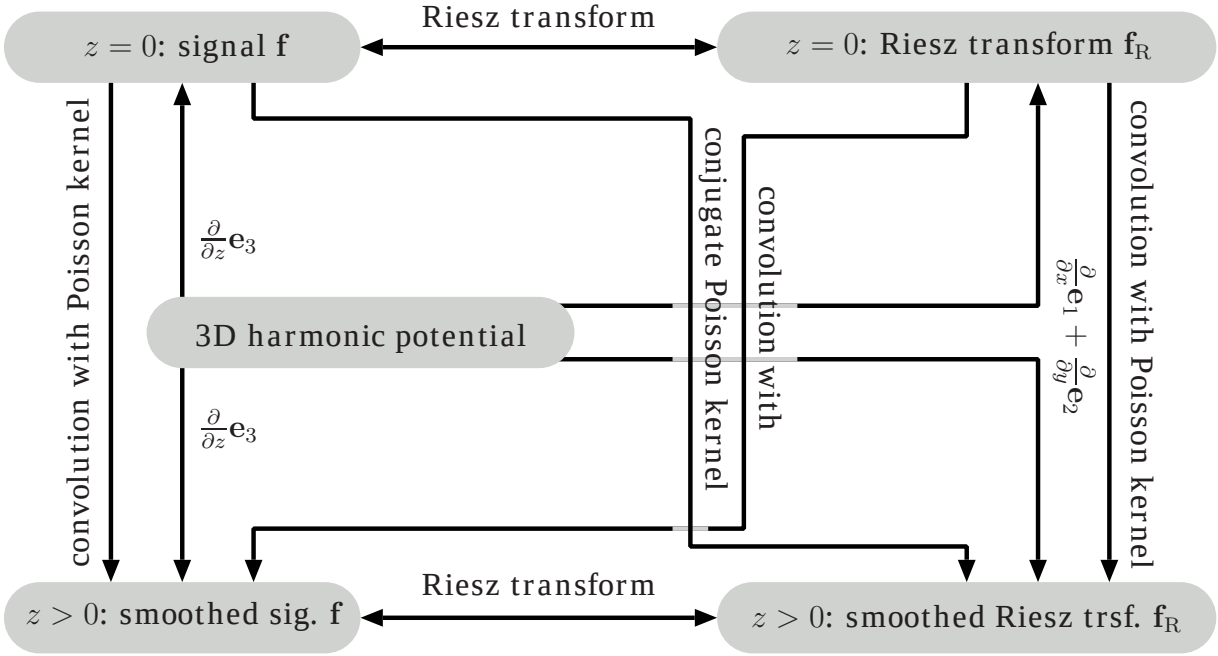

Fig. 4.4: Relationships between harmonic potential, signal, Riesz transform, and convolution with the Poisson kernel.

Riesz transform	$h_2(\mathbf{x}) = \frac{\mathbf{x}\mathbf{e}_3}{2\pi\lvert\mathbf{x}\rvert^3}$ (4.17) $H_2(\mathbf{u}) = \frac{\mathbf{u}}{\lvert\mathbf{u}\rvert} I_2^{-1}$ (4.18)
Poisson kernel (lowpass)	$l_s(\mathbf{x}) = \frac{s}{2\pi\lvert\mathbf{x} + s\mathbf{e}_3\rvert^3}$ (4.19) $L_s(\mathbf{u}) = \exp(-2\pi\lvert\mathbf{u}\rvert s)$ (4.20)
conjugate Poisson kernel	$h_s(\mathbf{x}) = \frac{\mathbf{x}\mathbf{e}_3}{2\pi\lvert\mathbf{x} + s\mathbf{e}_3\rvert^3}$ (4.21) $H_s(\mathbf{u}) = \frac{\mathbf{u}}{\lvert\mathbf{u}\rvert} I_2^{-1} \exp(-2\pi\lvert\mathbf{u}\rvert s)$ (4.22)

Tab. 4.1: Overview of the results about the Riesz transform, the Poisson (lowpass) filter, and the conjugate Poisson filter. These notations are used in the following sections and chapters.

By setting $p_s(\mathbf{x}) = k(\mathbf{x})L(s)$, (4.23) yields

$$\Delta_2 k(\mathbf{x})L(s) = -\frac{\partial^2}{\partial s^2}k(\mathbf{x})L(s)$$

and by dividing both sides by $L(s)$ and $k(\mathbf{x})$, two separate inhomogeneous equations are obtained:

$$\frac{\Delta_2 k(\mathbf{x})}{k(\mathbf{x})} = -\lambda^2 = -\frac{\frac{\partial^2}{\partial s^2}L(s)}{L(s)} \quad .$$

Setting $\lambda^2 = 4\pi^2|\mathbf{u}|^2$ where $\mathbf{u} = u\mathbf{e}_1 + v\mathbf{e}_2$ is a new vector variable, these two equations are solved by

$$k(\mathbf{x}, \mathbf{u}) = \exp(\pm I_3 2\pi\, \mathbf{x} \cdot \mathbf{u}) \tag{4.24}$$

and

$$L(\mathbf{u}, s) = \exp(\pm 2\pi|\mathbf{u}|s) \quad . \tag{4.25}$$

Hence, the Fourier kernel $k(\mathbf{x}, \mathbf{u})$ and the frequency response of the Poisson kernel are obtained (up to the sign). This way of solving the Laplace equation is closely related to the solution of the heat equation in [95] such that the relationship to the Gaussian scale-space becomes evident (see also section 4.5).

The Riesz transform is not well known in engineering, although its properties are quite attractive for applications. In mathematics however, the Riesz transform has been analyzed in detail in the field of Calderón-Zygmund theory. The aim of that theory was to investigate the multidimensional generalization of the Hilbert transform [139] and it was started in the fifties by Calderón and Zygmund at Chicago university. The first phase of this theory was concluded by the books of Stein [137], 1970, and Stein and Weiss [136], 1971, covering nearly all the here presented mathematics (although the relationship between nD singular operators and $(n+1)D$ harmonic functions is better reflected in [138]). A more recent work which also reflects the relationship to wavelet theory is [100].

In geophysics the Riesz transform was already used in 1984 [109], where the relation to harmonic functions was used to compute the earth's magnetic field from its partial derivatives on the surface.

In optics several authors started using the Riesz transform independently at the end of the nineties:

- Davis et al. applied the Riesz transform for detecting edges [32]. Since they did not use the term 'Riesz transform', I suppose they were not aware of it. Looking solely at the Riesz transform, they make no use of a phase concept as it is done in the subsequent section.
- Müller and Marquard use the Riesz transform to form a generalized analytic signal already in 1999 [108]. However, their results differ totally from those

presented in the following section, because they combine the two parts of the Riesz transform by addition and multiplication according to adding and multiplying the partial Hilbert transforms as proposed in [66].

- Larkin et al. use the Riesz transform combined with the original signal to demodulate fringe patterns [93, 92]. Although the idea of their approach is close to some aspects which are presented in this chapter, they investigate the Riesz transform mainly with respect to the specific application of interferograms (see also [20, 112]).

Although the mentioned optic papers are all related to image processing, none of them establishes the Riesz transform as a general tool for pattern recognition superior to the classical approach of quadrature filters.

Not being aware of Calderón-Zygmund theory nor the papers [109, 108], I derived the Riesz transform independently from geometric constraints [45, 50]. Due to the references [136, 109] which I got from my former colleague T. Bülow I was able to identify the derived transform with the Riesz transform. The study of further references about singular operators, harmonic and Clifford analysis finally led to the current state of the presented work.

4.3 The Monogenic Signal

The 1D analytic signal f_A can be considered as the holomorphic extension of a signal f restricted to the line $y_0 = 0$. Therefore, one can conclude that the term 'analytic signal' is related to 'analytic function' (which is equivalent to 'holomorphic function', see footnote 10 on page 32). In the previous section, I have shown that the monogenic extension of a 2D signal f is uniquely given by means of the Poisson kernel and its conjugate kernel. Hence, restricting this 3D monogenic function to the plane $s = 0$ yields the 2D generalization of the analytic signal. According to the discussion about the term 'analytic', this generalized 2D analytic signal is called the *monogenic signal*.

4.3.1 Definition and Properties

To be more explicit, the monogenic signal is defined by

$$\mathbf{f}_\mathrm{M}(\mathbf{x}) = \mathbf{f}(\mathbf{x}) + \mathbf{f}_\mathrm{R}(\mathbf{x}) \ , \tag{4.26}$$

where

$$\mathbf{f}_\mathrm{R}(\mathbf{x}) = (h_2 * \mathbf{f})(\mathbf{x}) = \iint \frac{\mathbf{x}'\mathbf{e}_3}{2\pi|\mathbf{x}'|^3}\mathbf{f}(\mathbf{x}-\mathbf{x}')\,dx'\,dy' \tag{4.27}$$

is the Riesz transform of $\mathbf{f}$. In the frequency domain, these definitions are equivalent to

$$\mathbf{F}_{\mathrm{M}}(\mathbf{u}) = \mathbf{F}(\mathbf{u}) + \mathbf{F}_{\mathrm{R}}(\mathbf{u}) = (1 + H_2(\mathbf{u}))\mathbf{F}(\mathbf{u}) = \left(1 + \frac{\mathbf{u}}{|\mathbf{u}|} I_2^{-1}\right) \mathbf{F}(\mathbf{u}) \ . \tag{4.28}$$

Now, having defined the monogenic signal as a 2D generalization of the analytic signal, the second step is to check whether the properties of the latter are fulfilled. First some properties of the Riesz transform are considered:

- It is anti-symmetric since $H_2(-\mathbf{u}) = -H_2(\mathbf{u})$. Note in this context that symmetry in 2D can be with respect to a point or to a line. The choice of symmetry is fundamental to designing the generalization of the Hilbert transform (in 1D there exists only one symmetry). Obviously, the Riesz transform corresponds to point-symmetry, whereas the approach in [17] corresponds to line-symmetry with respect to the coordinate axes and the partial Hilbert transform corresponds to line-symmetry with respect to a preference direction.
- It suppresses the DC component. A singularity exists at $\mathbf{u} = 0$. If it is removed by continuously extending the two components of the Riesz transform along the lines $u = 0$ and $v = 0$ (see (4.15)), $H_2(0) = 0$ is immediately obtained.
- The amplitude response is of value one for all non-zero frequencies: $|H_2(\mathbf{u})| = 1$ $\forall \mathbf{u} \neq 0$. This follows directly from the definition of $H_2(\mathbf{u})$ in (4.15).

These properties can be verified by considering the vector field H_2 (see figure 4.5). The multiplication of $\mathbf{u}$ with I_2^{-1} from the right in (4.18) just yields the dual of $\mathbf{u}$ with respect to the plane $\mathbf{e}_1 \wedge \mathbf{e}_2$. The 2D dual of a vector is given by the orthogonal vector, or the $90°$ rotated vector in the $\mathbf{e}_1 \wedge \mathbf{e}_2$-plane.

According to the properties of the Riesz transform, and in comparison to the analytic signal, the monogenic signal has the following properties:

- Its energy is two times the energy of the original signal if the DC component is neglected:

$$\iint |\mathbf{f}_{\mathrm{M}}(\mathbf{x})|^2 \, dx \, dy = \iint \mathbf{f}(\mathbf{x})^2 + |\mathbf{f}_{\mathrm{R}}(\mathbf{x})|^2 \, dx \, dy = 2 \iint \mathbf{f}(\mathbf{x})^2 \, dx \, dy \ .$$

- In polar coordinates the monogenic signal fulfills the split of identity. Since the energy is only modified by a constant real factor, the direct conclusion is that the amplitude of the monogenic signal is *isotropic*, which means that there is no dependence on the orientation of a signal (see also figure 4.6). This can be compared to the isotropy property of the structure tensor, which fulfills the invariance – equivariance property with respect to energy and orientation (but not phase). Further details will be discussed in section 4.4.

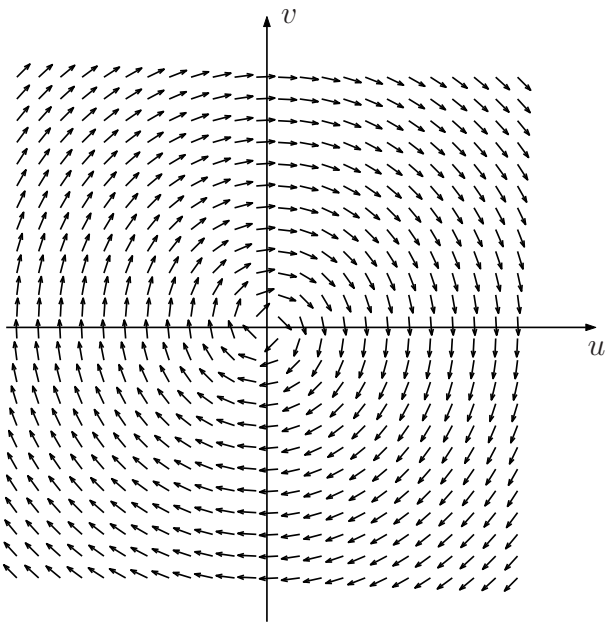

Fig. 4.5: Frequency response of the Riesz transform displayed as a 2D vector field. The axes are the frequencies u and v. The vector field is given by the 2D dual of the normalized frequency vectors (see (4.15)) and by setting it to zero at the origin (see text).

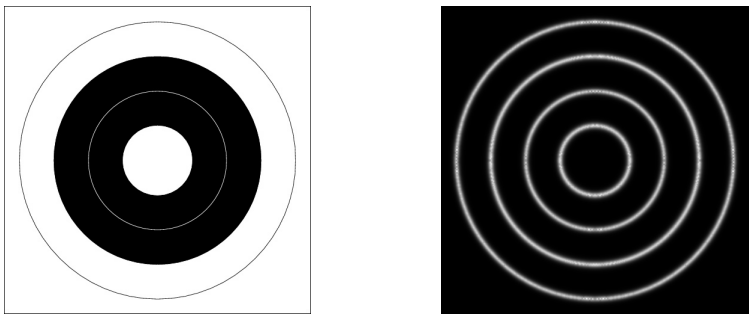

Fig. 4.6: Left: test image showing all even and odd symmetries for all orientations; right: the energy of the corresponding monogenic signal is isotropic and independent of the local symmetry.

- The Fourier transform of the monogenic signal is not one-sided, i.e., it includes redundancies, but this property is irrelevant for pattern recognition. Nevertheless, it is possible to use a non-redundant representation [45].

Not having a 'one-sided' spectrum is no fundamental drawback. In 1D the analytic signal is one-sided in the frequency domain, which is directly related to the prin-

ciple of causality of 1D signals by means of the Laplace transform (see (4.8)). In 2D however, signals are *intrinsically non-causal* (unless some time-dependency is introduced [128]), and hence it is not reasonable to demand for one-sidedness. Obviously, this results in losing the direct relation to the 2D Laplace transform which is based on a tensor approach (see e.g. [16], section 12.1.5).

However, due to the symmetry relation of the 2D Fourier transform for $\mathbf{e}_3$-valued signals (see (3.28)), the computation of the monogenic signal can be reduced to *any* half-space (see appendix A.4):

$$\mathbf{f}_{\mathrm{M}}(\mathbf{x}) = \langle 2 \int_0^{\infty} \int_{\theta_0}^{\theta_0+\pi} (1 + \mathbf{n}_\theta I_2^{-1}) \mathbf{F}(q\mathbf{n}_\theta) \exp(I_3 2\pi q \mathbf{n}_\theta \cdot \mathbf{x}) |q| \, d\theta \, dq \rangle_1$$

where in practical applications it is preferable to fix θ_0 to one of the coordinate axes. The summation in the case of the discrete Fourier transform can be performed for $u \geq 0$, where it is important to have $v > 0$ for $u = 0$.

4.3.2 The Definition of the Local Phase

As explained in section 3.2, the local amplitude and local phase of the 1D analytic signal are given by the logarithm of $\mathbb{R}_2^+$ (see (3.15) and (3.16)).

In the light of the definition of local phase and local amplitude in 1D, it is straightforward to use the logarithm of $\mathbb{R}_3^+$ for defining a 2D phase concept. The spinor field which maps the field $\mathbf{f}_{\mathrm{Id}}(\mathbf{x}) \equiv \mathbf{e}_3$ to $\mathbf{f}_{\mathrm{M}}(\mathbf{x})$ is given by $\mathbf{f}_{\mathrm{M}}(\mathbf{x})\mathbf{e}_3$. The logarithm of that spinor field yields the logarithm of the local amplitude

$$|\mathbf{f}_{\mathrm{M}}(\mathbf{x})| = \exp(\log(|\mathbf{f}_{\mathrm{M}}(\mathbf{x})|)) = \exp(\langle \log(\mathbf{f}_{\mathrm{M}}(\mathbf{x})\mathbf{e}_3) \rangle_0) \tag{4.29}$$

and the *local rotation vector*

$$\mathbf{r}(\mathbf{x}) = \arg(\mathbf{f}_{\mathrm{M}}(\mathbf{x}))^* = \langle \log(\mathbf{f}_{\mathrm{M}}(\mathbf{x})\mathbf{e}_3) \rangle_2^* \ . \tag{4.30}$$

This local rotation vector can be seen as the generalization of the local phase in 1D. At those positions where $\mathbf{f}_{\mathrm{M}}(\mathbf{x}) = -|\mathbf{f}_{\mathrm{M}}(\mathbf{x})|\mathbf{e}_3$, the local rotation vector is not well defined. Any vector lying in the plane $\mathbf{e}_1 \wedge \mathbf{e}_2$ with magnitude π is a correct solution.[6] For this case, the following extension of the definition can be made. The underlying vector field is supposed to be smooth. Considering the values of (4.30) in an open sphere with radius ε around the singular point and letting ε tend to zero while averaging the rotation axes, a well defined rotation vector is obtained. This interpolation of the rotation axis at singular points is used in [44] for a stable orientation estimation.

[6] In 1D this ambiguity reduces to the sign of the phase: $-\pi$ and π are both correct.

The rotation vector $\arg(\mathbf{f}_M)$ always lies in the plane orthogonal to $\mathbf{e}_3$ since it is dual to the wedge product $\mathbf{f}_M \wedge \mathbf{e}_3$ (see (4.30) and (2.12)). Hence, the local rotation vector field can be visualized by a 2D vector field. Comparable to phase wrapping in 1D (see also footnote 6), there is a *wrapping of the rotation vectors* of the 2D local phase: if a vector in a certain direction would exceed the amplitude π, it is replaced by the vector minus 2π times the unit vector in that direction, i.e., it points in the opposite direction, see figure 4.7.

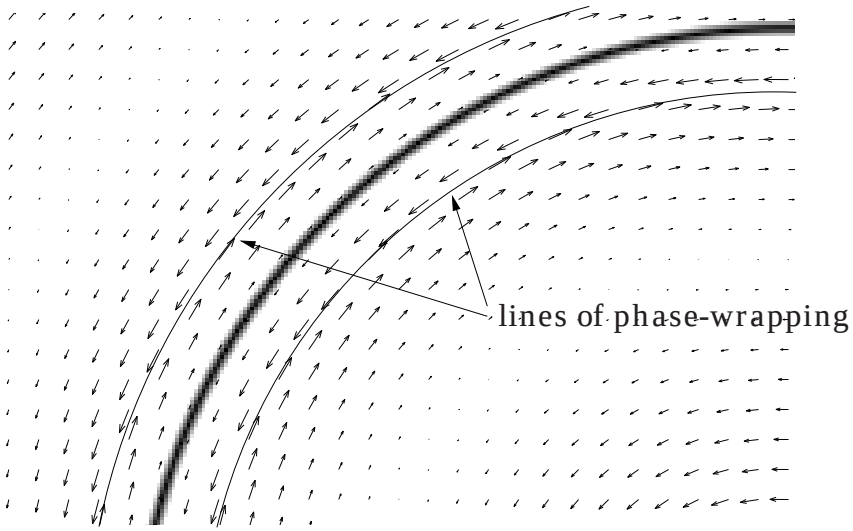

Fig. 4.7: Local phase vectors of one part of figure 4.6, left (curved line, displayed as underlying grayscale image). The phase is represented by the vector field which is obtained according to (4.30). Hence, the length indicates the rotation angle and the direction indicates the rotation axis. Note the phase-wrapping.

The choice of definitions for the local phase vector (4.30) and local amplitude (4.29) establishes a *transformation to spherical coordinates*. Given the local phase vector $\arg(\mathbf{f}_M)$ and the local amplitude $|\mathbf{f}_M|$ of a monogenic signal, it can be reconstructed by

$$\mathbf{f}_M(\mathbf{x}) = |\mathbf{f}_M(\mathbf{x})| \exp(\arg(\mathbf{f}_M(\mathbf{x}))) \ . \tag{4.31}$$

This directly follows from the proof about the $\mathbb{R}_3^+$-logarithm at the end of section 2.2.

As stated above, the monogenic signal fulfills indeed the split of identity since the local phase vector and the local amplitude are orthogonal. The local amplitude includes energetic information and the phase vector includes structural information, similar to the 1D analytic signal. In contrast to the 1D case, the phase now includes additional *geometric information*. The structural and geometric information are again orthogonal and therefore fulfill the invariance – equivariance property which is discussed in detail at the end of the next section.

4.4 The Interpretation of the Local Phase

Up to now, I have considered the 1D analytic signal and the 2D monogenic signal as two distinct approaches, which are only related by the fact that the latter is the generalization of the former with respect to the dimension of the domain. In signal theory there is a well known relation between 1D and 2D signals, the Radon transform [119]. The Radon transform can be used to transfer the 1D phase interpretation to the new 2D phase.

4.4.1 Transferring the 1D Phase to 2D

The Radon transform $\mathcal{R}$ maps a 2D signal onto an orientation parameterized family of 1D signals by integrating the 2D signal on the line given by the orientation parameter,

$$\mathcal{R}\{\mathbf{f}(\mathbf{x})\}(t,\theta) = \iint \mathbf{f}(\mathbf{x})\delta_0(\mathbf{x}\cdot\mathbf{n}_\theta - t)\,dx\,dy \ , \tag{4.32}$$

where $\delta_0(\cdot)$ is the Dirac delta and $\mathbf{n}_\theta = \cos\theta\mathbf{e}_1 + \sin\theta\mathbf{e}_2$ is the normal vector given by the orientation[7] $\theta \in [0,\pi)$. Geometrically, the Radon transform projects (orthogonally) the 2D signal onto a line with orientation θ.

The Radon transform is an invertible transform with several neat properties such as the Fourier slice theorem [74],

$$\mathcal{F}_1\{\mathcal{R}\{\mathbf{f}\}\}(u,\theta) = \mathbf{F}^{\mathrm{e}}(u\mathbf{n}_\theta) + \mathbf{e}_3\mathbf{F}^{\mathrm{o}}(u\mathbf{n}_\theta) \ . \tag{4.33}$$

The base vector $\mathbf{e}_3$ must be introduced for the odd part, because in the 1D Fourier transform I_2 is used, whereas in the 2D Fourier transform I_3 is used. The slice theorem says that the Radon transform projects the 2D signal in such a way, that it is *decomposed into its intrinsically 1D parts* (see also the discussion about the intrinsically 1D basis functions of the Fourier transform in section 3.4). The Fourier theory ensures that this is done in an optimal way with respect to the L^2-norm.

As a consequence of the slice theorem, the Radon transform relates the Riesz transform of a 2D signal to the Hilbert transforms of the 1D signals obtained from the Radon transform.[8] Accordingly, one can conclude that the interpretation of the monogenic phase vector of intrinsically 1D signals in 2D space is basically the same as the interpretation of the complex phase in 1D space. However, also for intrinsically 2D signals an interpretation is obtained by decomposing the signal into its

[7] Note the difference between *direction* and *orientation* in this context: a direction corresponds to a vector, an orientation to a 1D subspace.

[8] Actually, this relationship is well covered by the results from Calderón-Zygmund theory [139], which is not very accessible to engineers and computer scientists due to its abstract mathematical formulation. Therefore, a calculus-based derivation of these results is surely helpful.

intrinsically 1D parts. Indeed, the Radon transform is the connecting link between the 1D and 2D approaches.

The Radon transform of the Riesz transform $\mathbf{f}_{\mathrm{R}}(\mathbf{x})$ of an $\mathbf{e}_3$-valued 2D signal $\mathbf{f}(\mathbf{x})$ is given by the Hilbert transform of the Radon transform of $\mathbf{f}(\mathbf{x})$ according to

$$\mathcal{R}\{\mathbf{f}_{\mathrm{R}}(\mathbf{x})\}(t,\theta) = h_1(t\mathbf{e}_1) * \mathcal{R}\{\mathbf{f}(\mathbf{x})\}(t,\theta)\mathbf{n}_\theta^* \ . \tag{4.34}$$

This equivalence can easily be shown in the Fourier domain using the linearity of the Radon, Riesz, Hilbert, and Fourier transforms:

$$\begin{aligned}
\mathcal{F}_1\{\mathcal{R}\{\mathbf{f}_{\mathrm{R}}\}\}(u,\theta) &\overset{(4.33)}{=} \mathbf{F}_R^{\mathrm{e}}(u\mathbf{n}_\theta) + \mathbf{e}_3\mathbf{F}_R^{\mathrm{o}}(u\mathbf{n}_\theta) \\
&= \frac{u\mathbf{n}_\theta}{|u\mathbf{n}_\theta|} I_2^{-1}\mathbf{F}^{\mathrm{o}}(u\mathbf{n}_\theta) + \mathbf{e}_3\frac{u\mathbf{n}_\theta}{|u\mathbf{n}_\theta|} I_2^{-1}\mathbf{F}^{\mathrm{e}}(u\mathbf{n}_\theta) \\
&= \mathrm{sign}(u)\mathbf{n}_\theta I_2^{-1}(\mathbf{F}^{\mathrm{o}}(u\mathbf{n}_\theta) - \mathbf{e}_3\mathbf{F}^{\mathrm{e}}(u\mathbf{n}_\theta)) \\
&= -H_1(u\mathbf{e}_1)(\mathbf{F}^{\mathrm{o}}(u\mathbf{n}_\theta) + \mathbf{e}_3\mathbf{F}^{\mathrm{e}}(u\mathbf{n}_\theta))\mathbf{n}_\theta I_2^{-1} \\
&= H_1(u\mathbf{e}_1)(\mathbf{F}^{\mathrm{e}}(u\mathbf{n}_\theta) + \mathbf{e}_3\mathbf{F}^{\mathrm{o}}(u\mathbf{n}_\theta))\mathbf{n}_\theta^* \\
&= \mathcal{F}_1\{h_1 * \mathcal{R}\{\mathbf{f}\}\}(u,\theta)\mathbf{n}_\theta^* \ .
\end{aligned}$$

The Hilbert transform of the $\mathbf{e}_3$-valued function $\mathbf{f}$ is I_3 valued, and therefore the multiplication with the dual normal vector yields a vector parallel to $\mathbf{n}_\theta$. Hence, this theorem says that the Riesz transform is given by the *linear superposition of the oriented Hilbert transforms of the intrinsically 1D decomposition of the signal* (see also the discussion of the slice theorem). Thus, the Radon transform provides a way to calculate the Riesz transform (and therefore also the monogenic signal) by using the Hilbert transform (see also figure 4.8). This relationship has several consequences, e.g. for the processing of X-ray tomography data and for the design of quadrature filters [46, 47].[9]

The most important consequence of the previous relationship is that *the interpretation of the monogenic phase vector is given directly by the phase of the analytic signal*, since

$$\begin{aligned}
\mathcal{R}\{\mathbf{f}_{\mathrm{M}}\}(t,\theta) &= \mathcal{R}\{\mathbf{f}\}(t,\theta) + \mathcal{R}\{\mathbf{f}_{\mathrm{R}}\}(t,\theta) \\
&= \mathcal{R}\{\mathbf{f}\}(t,\theta) + h_1(t) * \mathcal{R}\{\mathbf{f}\}(t,\theta)\mathbf{n}_\theta^* \\
&= \tilde{s}_\theta(\mathcal{R}\{\mathbf{f}\}(t,\theta) + h_1 * \mathcal{R}\{\mathbf{f}\}(t,\theta)\mathbf{e}_1^*)s_\theta \ ,
\end{aligned}$$

where the spinor $s_\theta = \exp(\mathbf{e}_{12}\theta/2)$ represents the rotation about the $\mathbf{e}_3$-axis by θ. Evaluating the last line for $\theta = 0$ gives $\mathcal{R}\{\mathbf{f}\}(t,\theta) + h_1(t) * \mathcal{R}\{\mathbf{f}\}(t,\theta)\mathbf{e}_1^*$, i.e., basically the analytic signal of $\mathcal{R}\{\mathbf{f}\}(t,\theta)$. For any $\theta = \theta_0$ the Radon transform of the monogenic signal is just the analytic signal of the Radon transform of the signal, but with the base vector $\mathbf{e}_1$ rotated by θ_0: $\tilde{s}_{\theta_0}\mathbf{e}_1 s_{\theta_0} = \cos\theta_0\mathbf{e}_1 + \sin\theta_0\mathbf{e}_2$. The impact of this

[9] In this thesis I use a different method for designing quadrature filters.

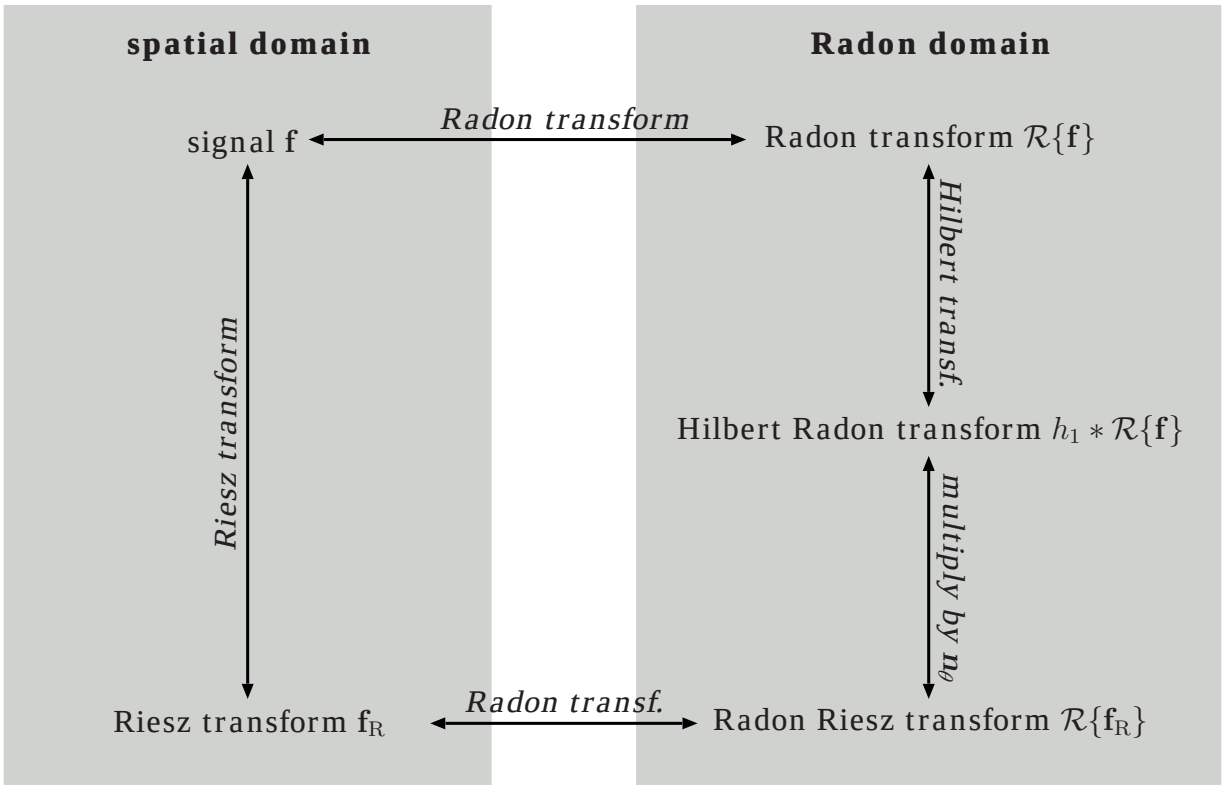

Fig. 4.8: The relation between Radon, Riesz, and Hilbert transform. The Riesz transform in the spatial domain (left) is equivalent to the Hilbert transform with subsequent multiplication by the orientation vector in the Radon domain (right).

result on the interpretation of the monogenic phase vector is that for linear structures with large support (lines, edges in images), the Radon transform is dominated by this structure. Hence, the monogenic phase is given by the main orientation θ_0 and the 1D phase of $\mathcal{R}\{\mathbf{f}\}(t, \theta_0)$. In the idealized case (i.e., the signal is constant in one direction), the Radon transform is non-zero only for the main orientation. Accordingly, the $\mathbf{n}_\theta$ in (4.34) is constant (equals $\mathbf{n}_{\theta_0}$) and the Riesz transform is given by

$$\mathbf{f}_\mathrm{R}(\mathbf{x}) = (\delta_0(\mathbf{x} \wedge \mathbf{n}_{\theta_0}) h_1(\mathbf{x} \cdot \mathbf{n}_{\theta_0})) * \mathbf{f}(\mathbf{x}) \mathbf{n}^*_{\theta_0} \quad , \tag{4.35}$$

which is just the partial Hilbert transform of $\mathbf{f}(\mathbf{x})$ with preference direction θ_0 multiplied by $\mathbf{n}^*_{\theta_0}$. Thus, (4.34) is simplified to theorem 1 in [50]. In contrast to the latter theorem, (4.34) also provides a reasonable interpretation for non-perfect i1D signals.

4.4.2 Decomposition of the Phase Vector

For an intrinsically 1D signal the following decomposition of the phase vector (4.30) is obtained while keeping the local phase of the underlying 1D signal: the rotation vector $\mathbf{r} = \arg(\mathbf{f}_M(\mathbf{x}))$ is orthogonal to the *local orientation* (i.e., the orientation of the Dirac line in the Fourier domain) of the 2D signal and $\operatorname{sign}(\mathbf{r} \cdot \mathbf{e}_1)|\mathbf{r}|$ represents the *local i1D phase* of the 2D signal (see also figure 4.9). These interpretations are consistent with the former definition of local phase and local orientation in [45].

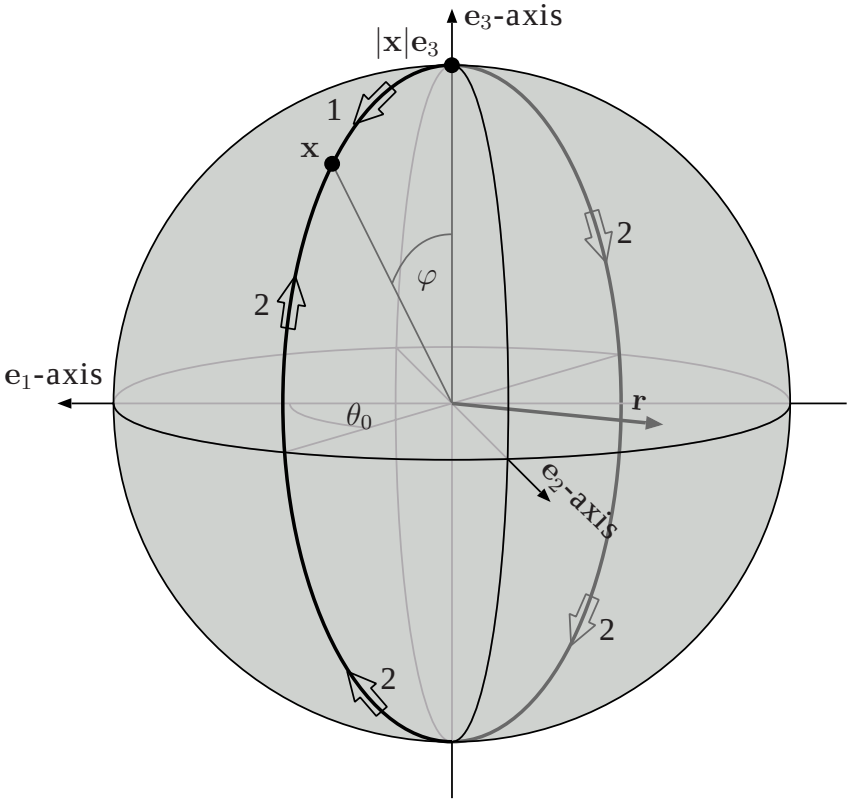

Fig. 4.9: Phase representation using a rotation vector $\mathbf{r}$. First the vector $|\mathbf{x}|\mathbf{e}_3$ is rotated about the $\mathbf{e}_2$-axis by φ, and second about the $\mathbf{e}_3$-axis by θ_0. The rotation vector is orthogonal to the plane spanned by the $\mathbf{e}_3$-axis and the vector $\mathbf{x}$. Its length is given by the angle between the $\mathbf{e}_3$-axis and $\mathbf{x}$. Although the rotation vector is unique, the decomposition is not. There are two possibilities: 1. (θ_0, φ) and 2. $(\theta_0 + \pi, -\varphi)$.

Instead of considering the local rotation vector as the 2D equivalent of the 1D local phase, one can also think of a generalized complex phase which might be easier to understand for people working on classical complex signal theory. Since the argument of the monogenic signal is defined by the bivector part of the logarithm

of $\mathbf{f}_M\mathbf{e}_3$ (see (4.30)), the interpretation has changed from a vector valued signal to an $\mathbb{R}_3^+$-valued signal. The even subalgebra of $\mathbb{R}_3$ is isomorphic to the algebra of quaternions. Furthermore, any unit quaternion with zero real part (in $\mathbb{R}_3$: any unit bivector B) can substitute the imaginary unit of the complex numbers (see section 2.2). In the context of the phase interpretation of the monogenic signal, one can think of $\{1, B\}$ being the complex plane that is given by the local orientation, i.e., defined by 1, $\mathbf{f}_M\mathbf{e}_3$, and the origin (see figure 4.9: the great circle trough $|\mathbf{x}|\mathbf{e}_3$ and $\mathbf{x}$ is the intersection of the sphere with radius $|\mathbf{x}|$ and this plane). The local phase is then obtained by the ordinary complex argument in this plane. Hence, the interpretation of the local rotation vector has changed to a complex phase in the complex plane where the imaginary unit is given by the orientation vector.

Unfortunately, both of these definitions do not yield a unique phase representation, since a rotation of the signal by π yields the same orientation and a negated phase. This ambiguity can be visualized by two different decompositions of the rotation vector (see figure 4.9). The signal theoretic background of this ambiguity is the following. From a local signal it is only possible to recover the local orientation, the directional sense is lost. A typical example is an i1D signal as in figure 3.6 where it is unknown whether the direction of the signal is parallel or anti-parallel to $\mathbf{n}$. The same problem also occurs in the context of oriented quadrature filters (see [61]), where Granlund and Knutsson claim that there is no local way to get the direction from the orientation. For the evaluation of the local phase one has to fix the local direction either to be equal to the local orientation (i.e., in $[0, \pi)$) or to be opposite to it (i.e., in $[\pi, 2\pi)$). If the wrong directional sense is chosen, the phase is inverted, and hence the local phase and the local direction show discontinuities. Unfortunately, the choice of the directional sense is arbitrary unless further *constraints* are introduced (see below).

The mathematical reason for the ambiguity in the phase-direction decomposition is that the group which emerges from phase and direction changes is not simply connected. A group is simply connected if it is connected, i.e., for any two group elements there exists a smooth path which connects them [69], and if every closed path can be reduced to a point [56]. If the direction and the phase are considered in the interval $[0, 2\pi) \times [0, 2\pi)$, the orbit of the corresponding group is a torus, the *direction-phase torus* (see figure 4.10). A torus is not simply connected [56], section 13.3b, and therefore two points on the torus do not uniquely determine their relationship. However, the 2D sphere which is created by considering the phase-orientation decomposition is simply connected [69], and therefore there exists no unique mapping from phase-orientation space to phase-direction space.

In order to create the complete direction-phase torus, the local rotation vectors (i.e., points on the 3D unit sphere) are first mapped onto one half-torus. Without loss of generality one can bijectively map the unit sphere to the outer half-torus[10]

[10] To be more exact: any open neighborhood of the unit sphere not containing the north-pole nor

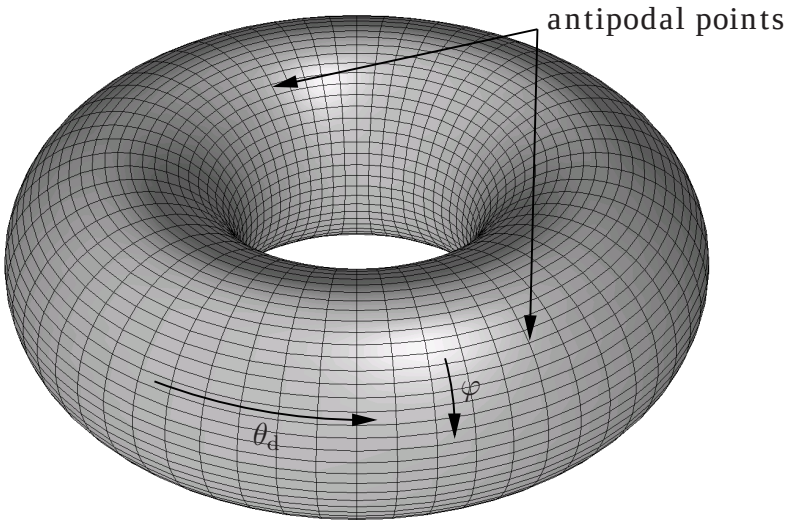

Fig. 4.10: Complete phase-direction model for 2D signals. The angle θ_d indicates the local direction and φ indicates the local phase. An orientation-phase pair corresponds to two antipodal points on the torus, one corresponds to case 1 and the other one corresponds to case 2 of figure 4.9.

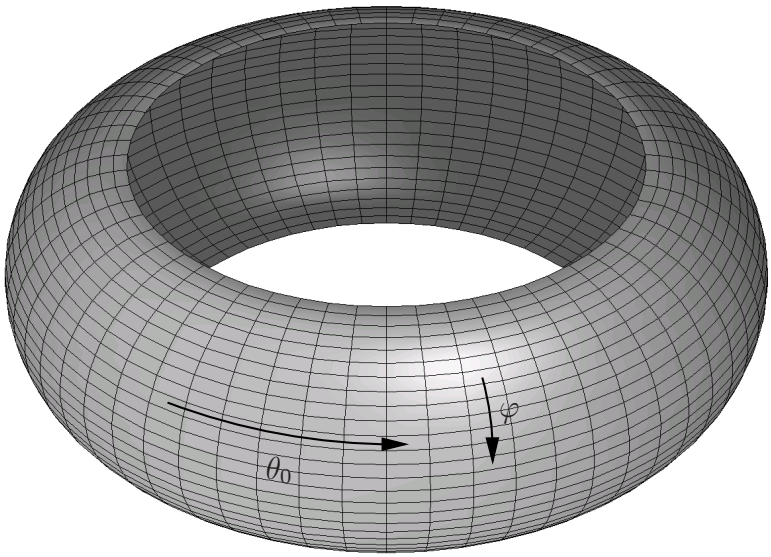

Fig. 4.11: Points of the unit sphere are mapped onto the outer half-torus. The angle θ_0 indicates either the local direction or the opposite direction. The angle φ indicates the local phase or the negated local phase.

(see figure 4.11) by the formula

$$\mathbf{x}_{\mathrm{T}} = R\frac{(\mathbf{x} \wedge \mathbf{e}_3) \cdot \mathbf{e}_3}{|\mathbf{x} \wedge \mathbf{e}_3|} + \mathbf{x} \ , \tag{4.36}$$

where $\mathbf{x}$ is a point on the sphere and $\mathbf{x}_{\mathrm{T}}$ is the corresponding point on the outer half-torus with great circle radius R. The two singular points $\mathbf{x} = |\mathbf{x}|\mathbf{e}_3$ and $\mathbf{x} = -|\mathbf{x}|\mathbf{e}_3$ are mapped onto the two boundary circles of the half-torus, which illustrates the orientation-singularity explained in section 4.3.2.

If the external constraints require to switch the directional sense and the sign of the phase, the position on the outer half-torus must be changed to the corresponding position on the inner half-torus. This corresponding position is obtained by the two-fold antipodal position (antipodal with respect to both circles of the torus, see figure 4.12). The first antipodal (with respect to the main circle) yields another point

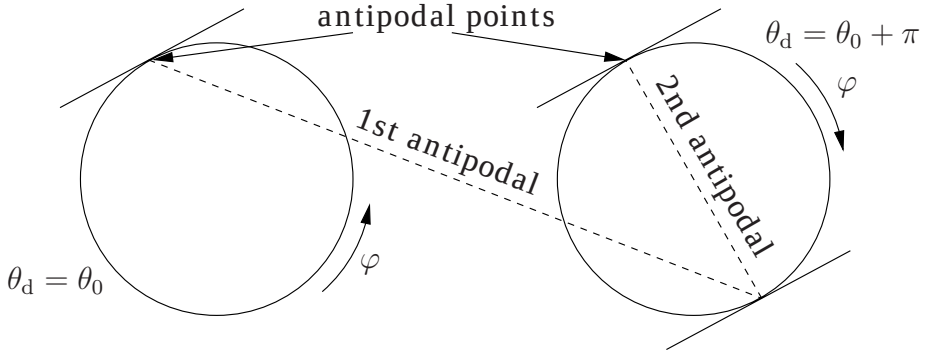

Fig. 4.12: Torus construction for a fixed local orientation θ_0. One orientation corresponds to two opposite directions $\theta_{\mathrm{d}} = \theta_0$ and $\theta_{\mathrm{d}} = \theta_0 + \pi$. The corresponding rotation planes (complex planes) also have opposite directional sense, indicated by the rotation sense of φ. Two antipodal points correspond to one orientation-phase pair. The antipodal with respect to the main circle and the minor circle is indicated in the figure. For the 3D position of the antipodals see also figure 4.10.

on the outer half-torus, whereas the subsequent second antipodal (with respect to the minor circle) yields the point on the inner half-torus. The two corresponding positions on the torus obviously correspond to the same rotation vector. The two-fold antipodal construction which is illustrated in figure 4.12 is obtained by the formula

$$\mathbf{x}_{\mathrm{T}} = -R\frac{(\mathbf{x} \wedge \mathbf{e}_3) \cdot \mathbf{e}_3}{|\mathbf{x} \wedge \mathbf{e}_3|} + \mathbf{x} \ , \tag{4.37}$$

the south-pole can be bijectively mapped to the outer half torus. The two poles are singular.

where $\mathbf{x}$ is again the point on the unit sphere and $\mathbf{x}_{\mathrm{T}}$ is the corresponding point on the inner half-torus. The only difference to (4.36) is the sign in front of R. The external constraints mentioned above just have to toggle this sign in order to obtain a full-torus, and hence a complete direction-phase decomposition.

The external constraints emerge from the application in which the phase is needed. The most obvious constraint is the *intra-image constraint* which is based on the fact that there is an upper bound of the *relative change of direction* between two local neighborhoods which is proportional to their distance in the image plane. The local direction function is continuous in the image plane, and therefore a small change in the image coordinates yields a bounded change of direction. In practice one can assume that there is a maximal gradient of the directional information. The question is, however, how to weight the angular distance of directions depending on the local phase (see also section 6.1). The easiest case is obviously a constant weight, which means that the directional sense is changed if the angular distance to the adjacent positions is greater than $\pi/2$. An example for this method is given in [93] and in [20]. In both papers it is proposed to apply an unwrapping of the orientation (modulo 2π) in order to obtain a consistent phase representation. The application in both papers is the 2D demodulation needed for interferogram processing.

If more than one image of the same physical object from different, but similar views is given (e.g. stereo images, image sequence) corresponding points in the images can be used to apply *inter-image constraints*. The idea of these constraints is that the local phase and the local direction as functions of the change of view are smooth to some degree. Hence, the direction-phase decompositions of the same physical point in different views are chosen to be compatible. An application using this constraint is given in section 6.3.

Concluding the discussion about the direction-phase decomposition, it is important to realize that there are always two solutions for consistent decompositions. If every point of a consistent solution is changed into its antipodal point, the direction-phase decomposition is consistent again. There is no way to prefer one of both solutions except for having external information.

Finally, if the monogenic phase is decomposed into local direction and local phase, the split of identity (the third property of the analytic signal) is also preserved with respect to geometric and structural information. The local phase is invariant to changes of the local direction and the local direction is invariant to changes of the local structure. For the pair orientation and phase, the split of identity is only fulfilled up to the ambiguity explained above.

4.5 A New Scale-Space Approach

In this section the new scale-space which is generated by the Poisson kernel (see (4.1) and (4.9)) is considered more in detail. I will show that the Poisson kernel establishes a linear scale-space similar to the Gaussian scale-space. Based on this scale-space a multi-band decomposition is presented which leads to generalized quadrature filters.

4.5.1 The Scale-Space Axiomatic of Iijima

In [144] the theory of scale-space is claimed to have its roots in Japan, based on a set of five axioms proposed by Iijima in 1959. According to [144], Iijima proves in [73][11] that the Gaussian scale-space is the *unique solution* for the following set of axioms:

1. The generating kernel of a linear scale-space is *linear*.
2. The kernel is *shift invariant*.
3. The kernel fulfills the *semigroup property*.
4. The kernel is *scale invariant*.
5. The kernel *preserves positivity*.

The first two axioms are equivalent to the fact that the kernel is a convolution kernel or an LSI operator. The third axiom says that successive application of two operators with parameters s_1 and s_2 is equivalent to the application of the operator with parameter $s_3 = m(s_1, s_2)$ where m is a function representing the group multiplication[12]. Scale invariance means that scale transformations and the kernel commute which induces that the application of the kernel to a scaled image is equivalent to the output of the scaled operator (where the Jacobian must be taken into account). Preserving of the positivity is given if the operator output of any positive signal is again positive.

Checking the properties for the 1D and 2D Gaussian kernels (denoted by $g_\sigma(\mathbf{x})$) and Poisson kernels, respectively, results that all five axioms are fulfilled:

1. The Gaussian filter and the Poisson filter are both LSI operators.
2. See above.

[11] Since I do not know the Japanese language, I was not able to verify this reference in detail.
[12] In group theory the multiplication is defined by the concatenation of group actions (see e.g. [125]).

3. This axiom is easily verified in the frequency domain. For the Gaussian kernel (1D and 2D):

   $$G_{\sigma_1}(\mathbf{u})G_{\sigma_2}(\mathbf{u}) = \exp(-2\pi^2\mathbf{u}^2\sigma_1^2)\exp(-2\pi^2\mathbf{u}^2\sigma_2^2) = \exp(-2\pi^2\mathbf{u}^2(\sigma_1^2+\sigma_2^2)) = G_{\sigma_3}(\mathbf{u}) \ ,$$

   hence $\sigma_3 = m(\sigma_1,\sigma_2) = \sqrt{\sigma_1^2+\sigma_2^2}$. For the Poisson kernel (1D and 2D):[13]

   $$L_{s_1}(\mathbf{u})L_{s_2}(\mathbf{u}) = \exp(-2\pi|\mathbf{u}|s_1)\exp(-2\pi|\mathbf{u}|s_2) = \exp(-2\pi|\mathbf{u}|(s_1+s_2)) = L_{s_1+s_2}(\mathbf{u}) \ ,$$

   hence $m(s_1,s_2) = s_1+s_2$.

4. The scale invariance is checked in the spatial domain ($\lambda \in \mathbb{R}^+$). For the Gaussian kernel (1D):

   $$g_\sigma(\lambda\mathbf{x}) = \frac{1}{\sqrt{2\pi}\sigma}\exp\left(-\frac{(\lambda x)^2}{2\pi\sigma^2}\right) = \frac{1}{\lambda}\,\frac{1}{\sqrt{2\pi}\frac{\sigma}{\lambda}}\exp\left(-\frac{x^2}{2\pi\left(\frac{\sigma}{\lambda}\right)^2}\right) = \frac{1}{\lambda}\,g_{\frac{\sigma}{\lambda}}(\mathbf{x}) \ ,$$

   where λ^{-1} is just the inverse of the Jacobian. It is necessary to introduce the inverse Jacobian since the scale invariance is defined by the observation transformation (convolution):

   $$\int \mathbf{f}\left(\frac{\mathbf{x}}{\lambda}\right) g_\sigma(\mathbf{x})\,d\mathbf{x} = \int \mathbf{f}(\mathbf{x}')g_\sigma(\lambda\mathbf{x}')\lambda\,d\mathbf{x}' = \int \mathbf{f}(\mathbf{x}')g_{\frac{\sigma}{\lambda}}(\mathbf{x}')\,d\mathbf{x}' \ .$$

   Similar for the Poisson kernel (1D):

   $$l_s(\lambda\mathbf{x}) = \frac{s}{\pi((\lambda x)^2+s^2)} = \frac{1}{\lambda}\,\frac{\frac{s}{\lambda}}{\pi\left(x^2+\left(\frac{s}{\lambda}\right)^2\right)} = \frac{1}{\lambda}\,l_{\frac{s}{\lambda}}(\mathbf{x}) \ .$$

   For the Gaussian kernel (2D):

   $$g_\sigma(\lambda\mathbf{x}) = \frac{1}{2\pi\sigma^2}\exp\left(-\frac{(\lambda\mathbf{x})^2}{2\pi\sigma^2}\right) = \frac{1}{\lambda^2}\,\frac{1}{2\pi\left(\frac{\sigma}{\lambda}\right)^2}\exp\left(-\frac{\mathbf{x}^2}{2\pi\left(\frac{\sigma}{\lambda}\right)^2}\right) = \frac{1}{\lambda^2}\,g_{\frac{\sigma}{\lambda}}(\mathbf{x}) \ ,$$

   where λ^{-2} is the inverse of the Jacobian. For the Poisson kernel (2D):

   $$l_s(\lambda\mathbf{x}) = \frac{s}{2\pi((\lambda\mathbf{x})^2+s^2)^{\frac{3}{2}}} = \frac{1}{\lambda^2}\,\frac{\frac{s}{\lambda}}{2\pi\left(\mathbf{x}^2+\left(\frac{s}{\lambda}\right)^2\right)^{\frac{3}{2}}} = \frac{1}{\lambda^2}\,l_{\frac{s}{\lambda}}(\mathbf{x}) \ .$$

5. Both kernels, the Gaussian kernel and the Poisson kernel, are positivity preserving since they are both positive for all $x \in \mathbb{R}$ (1D case) and $\mathbf{x} \in \mathbb{R}^2$ (2D case).

[13] For the sake of clarity, I use s instead of y for the Poisson kernel of 1D signals, hence s indicates the scale parameter in the 1D and 2D framework.

Hence, the proof of Iijima seems to contain an error with respect to the uniqueness. Besides the papers of Iijima, there are other publications claiming that the Gaussian kernel is the unique kernel which creates a linear scale-space. In [95], section 3.5, Lindeberg states that any LSI operator which performs a smoothing operation can be decomposed into four primitive functions: the Gaussian kernel, the truncated exponential function, the translation, and the rescaling. Imposing the semigroup property and the symmetry of the kernel, the author claims that the Gaussian kernel is the only remaining candidate. Unfortunately, he missed the obvious solution[14] of combining the two truncated exponential functions which yields the Poisson kernel.

Another uniqueness proof can be found in [3], which is based on five conditions. The most important one is the *monotonicity condition* which requires that the local maxima decrease and the local minima increase with the scale. This implies that no new extrema are created by the smoothing kernel when the scale becomes coarser. The proof for the necessity of the Gaussian kernel is obtained by considering a special generic signal which requires the Gaussian kernel to fulfill the monotonicity condition, and hence, the Poisson kernel fails to fulfill this condition. However, the described proof is only valid for the 1D scale-space and in 2D the non-creation of local maxima cannot be guaranteed [95]. On the other hand, the Poisson kernel fulfills the *maximum principle* [26] which means that the scale-space has *no local maxima or minima* for $s > 0$. This is obtained by a simple observation. First of all, the signal in scale-space is a harmonic function, since it is one component of a holomorphic function (1D signal means 2D embedding) or of a monogenic function (2D signal means 3D embedding). In either case the signal in scale-space fulfills the Laplace equation, and therefore no local extremum is possible, since one second derivative being greater than zero implies that another one must be less than zero. Hence, the extrema are located either on the line $s = 0$ or on the line in infinity $s \to \infty$. In the latter case however, the signal is reduced to its DC-component, i.e., it is constant and does not contain any extremum.

4.5.2 Comparison to Gaussian Scale-Space

Regardless if the various scale-space axiomatics are fulfilled or not, it is reasonable to compare the properties of both scale-spaces. Obviously, both kernels are *lowpass filters*. In figure 4.13 both filters can be compared in spatial and frequency domain for the 1D case. The scale parameter of the 1D Poisson kernel is obtained by setting the maximum equal to the maximum of the Gaussian kernel:

$$(\mathbf{f}_2 \cdot \mathbf{e}_2)(\mathbf{x} = 0) = \frac{1}{\pi s} = \frac{1}{\sqrt{2\pi}\sigma} \quad \Rightarrow \quad s = \frac{\sqrt{2}\sigma}{\sqrt{\pi}} \ .$$

[14] Lindeberg uses the absolute value of the scale parameter in the truncated exponential functions to underline that the scale parameter is in $\mathbb{R}^+$. The semi-group property is obviously lost if a modulus operation is involved, but it is unnecessary in this case.

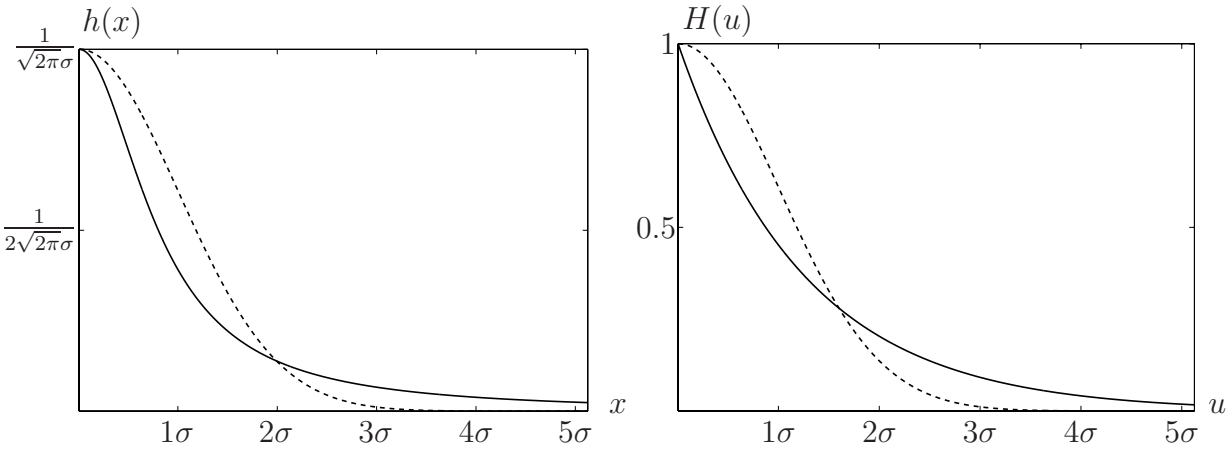

Fig. 4.13: The 1D Gaussian kernel (dashed line) and the 1D Poisson kernel (solid line) in spatial domain (left) and frequency domain (right). All axes are labeled relative to the standard deviation σ of the Gaussian.

Using this identity, the Gaussian kernel can be compared with the Poisson kernel. Obviously, the maximum of the Poisson kernel is sharper than that of the Gaussian kernel. However, the spatial extent of the Poisson kernel is larger than that of the Gaussian kernel. The same is true for the frequency response of the Poisson kernel. In practice the larger spatial extent yields either larger filter masks or an increased truncation error.

Besides this qualitative discussion, it is also possible to calculate the *uncertainties* of both kernels (for the formal definition of uncertainty see appendix A.3). The uncertainty of the Gaussian kernel is known to be the optimum $(4\pi)^{-1}$ (see (3.12) and [11]). The uncertainty of the Poisson kernel for 1D signals is slightly worse than that of the Gaussian kernel (factor $\sqrt{2}$, see appendix A.3).

In figure 4.14 both filters can be compared in spatial and frequency domain for the 2D case. Since the Gaussian filter looks the same in both domains, the frequency

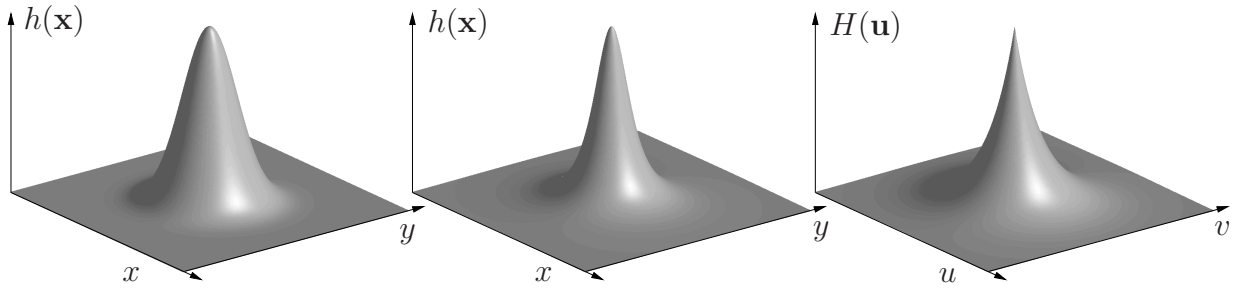

Fig. 4.14: The 2D Gaussian kernel (left) and the 2D Poisson kernel in spatial domain (middle) and frequency domain (right).

domain representation is omitted. The 2D Poisson filter, however, differs in the two domains. Again, the scale parameter of the Poisson kernel is obtained by setting the maxima of Gaussian kernel and Poisson kernel equal:

$$(\mathbf{f}_3 \cdot \mathbf{e}_3)(\mathbf{x} = 0) = \frac{1}{2\pi s^2} = \frac{1}{2\pi\sigma^2} \quad \Rightarrow \quad s = \sigma \ .$$

As in the 1D case, the peak of the Poisson kernel is sharper than that of the Gaussian kernel and the extent is larger.[15] The uncertainty of the 2D Poisson is only slightly worse than that of the Gaussian (factor $\sqrt{1.5}$, see appendix A.3).

In practice the behavior of both linear scale-spaces is nearly identical, which can be verified by successive application of both filters to an arbitrary image. At [41] mpeg-movies showing the descent through the scale-space can be found.

4.6 The Spherical Quadrature Filter

In this section the introduced scale-space is used to define a new kind of quadrature filter by means of the monogenic signal. The new quadrature filter is based on a radial bandpass filter which is obtained by taking the difference of Poisson kernels at different scales.

4.6.1 The DOP Wavelet

Both lowpass filters, the Gaussian kernel and the Poisson kernel, establish a linear scale-space. As in the case of the Gaussian kernel, it is possible to build up a *difference of Poisson filters* space (DOP). The difference of two Poisson kernels yields a bandpass. Combining two lowpasses with a fixed ratio of scale parameters yields a family of bandpasses with constant relative bandwidth (see also [48]),

$$\begin{aligned} B_{s_0,\lambda,k}(\mathbf{u}) &\stackrel{\text{def}}{=} L_{s_0\lambda^k}(\mathbf{u}) - L_{s_0\lambda^{k-1}}(\mathbf{u}) \\ &= \exp(-2\pi|\mathbf{u}|s_0\lambda^k) - \exp(-2\pi|\mathbf{u}|s_0\lambda^{k-1}) \ , \end{aligned} \tag{4.38}$$

where $\lambda \in (0,1)$ indicates the relative bandwidth, s_0 indicates the coarsest scale[16], and $k \in \mathbb{N}$ indicates the bandpass number (for $\lambda = 0.5$ k is the octave number). These radial bandpass filters are similar to the lognormal filters which is visualized in figure 4.15. The DOP filter decreases faster for high radial frequencies, but has

[15] This cannot be observed in figure 4.14, but the radial plots are similar to those in figure 4.13, left.

[16] Bandpass filters are DC-free. However for decreasing scale, the non-zero passband moves closer and closer to the origin such that there are high coefficients arbitrary close to the origin. For finite, discrete signals, there exists a coarsest scale such that the DC-component is zero. The remaining coarser part of the signal is the DC-component and corresponds to a signal size of one pixel in the case of a multi-resolution analysis.

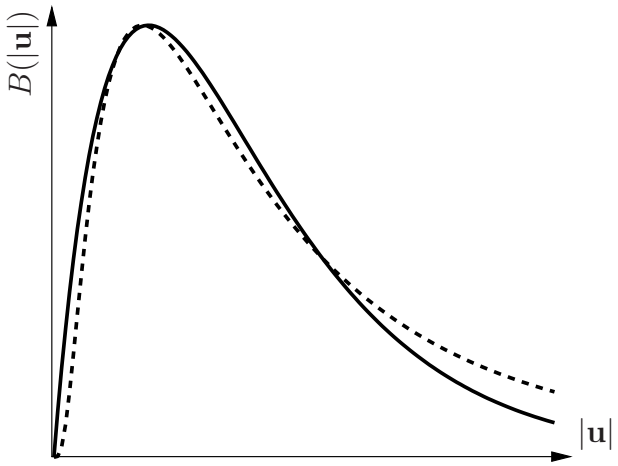

Fig. 4.15: DOP filter (solid line) and lognormal filter (dashed line) in radial coordinates.

a discontinuous first derivative in the origin. The bandpass filters defined by (4.38) are all self-similar. Two bandpasses with octave k_1 and k_2 are related by

$$B_{s_0,\lambda,k_1}(\mathbf{u}) = B_{s_0,\lambda,k_2}(\lambda^{k_1-k_2}\mathbf{u}) \ .$$

The corresponding convolution kernels are given by

$$b_{s_0,\lambda,k}(\mathbf{x}) = \frac{s_0\lambda^k}{2\pi(\mathbf{x}^2+(s_0\lambda^k)^2)^{3/2}} - \frac{s_0\lambda^{k-1}}{2\pi(\mathbf{x}^2+(s_0\lambda^{k-1})^2)^{3/2}} \tag{4.39}$$

and they yield a family of dilated LSI operators. By appropriate normalization, this family of bandpass filters can be used to create a continuous *wavelet transform*. The admissible condition of a 2D wavelet (see [96], 1.6.2.3) is fulfilled, since

$$\begin{aligned} C_B(s_1,s_2) &= \iint_{\mathbb{R}^2} \frac{|\exp(-2\pi|\mathbf{u}|s_1)-\exp(-2\pi|\mathbf{u}|s_2)|^2}{|\mathbf{u}|^2}\,du\,dv \\ &\quad \text{substitute } \mathbf{u} = q(\cos\theta\mathbf{e}_1+\sin\theta\mathbf{e}_2) \qquad (4.40) \\ &= 2\pi\int_{\mathbb{R}^+} \frac{\exp(-4\pi q s_1)-\exp(-2\pi q(s_1+s_2))+\exp(-4\pi q s_2)}{q}\,dq \\ &= 2\pi\log\left(\frac{s_1}{4s_2}+\frac{1}{2}+\frac{s_2}{4s_1}\right) \ . \qquad (4.41) \end{aligned}$$

For the case of $\lambda = 0.5$, the integral above yields

$$C_B(s_0\lambda^k, s_0\lambda^{k-1}) = 2\pi\log(\lambda/4+0.5+1/(4\lambda)) = 2\pi\log(9/8) \ .$$

Although $b_s(\mathbf{x})$ establishes a continuous wavelet transform, this does not automatically imply that $b_s(\mathbf{x})$ constitutes a *frame* or that it can be used for a *multiresolution analysis* (MRA). However, the DOP filters can be combined with subsampling operations in order to form a *scale-pyramid* according to the Laplace pyramid

(see [27]). This pyramid can be used to design efficient feature analysis algorithms. A straightforward application of this method is used in the disparity estimation algorithm in section 6.3. A further discussion on aspects of frame theory and MRA is out of the scope of this thesis.

4.6.2 The Definition of Spherical Quadrature Filter

The conjugate kernel which corresponds to $b_{s_0,\lambda,k}(\mathbf{x})$ is given by

$$c_{s_0,\lambda,k}(\mathbf{x}) = \frac{\mathbf{x}\mathbf{e}_3}{2\pi(\mathbf{x}^2 + (s_0\lambda^k)^2)^{3/2}} - \frac{\mathbf{x}\mathbf{e}_3}{2\pi(\mathbf{x}^2 + (s_0\lambda^{k-1})^2)^{3/2}} \tag{4.42}$$

with the frequency response

$$\begin{aligned} C_{s_0,\lambda,k}(\mathbf{u}) &= H_{s_0\lambda^k}(\mathbf{u}) - H_{s_0\lambda^{k-1}}(\mathbf{u}) \\ &= \frac{\mathbf{u}}{|\mathbf{u}|} I_2^{-1}(\exp(-2\pi|\mathbf{u}|s_0\lambda^k) - \exp(-2\pi|\mathbf{u}|s_0\lambda^{k-1})) \ . \end{aligned} \tag{4.43}$$

The triple consisting of the DOP filter and the *difference of conjugate Poisson filters* (DOCP) form a Riesz triple[17], i.e., the vector field formed by the filter output of the DOP and the DOCP is a harmonic field. In the light of the definition of quadrature filters (see section 3.2) as the combination of a bandpass filter with its Hilbert transform, the combined filter DOP+DOCP forms a generalized quadrature filter since it consists of a bandpass filter and its Riesz transform. Accordingly, the new filter is called a *spherical quadrature filter* (SQF). All the considerations about local phase, local orientation, and local amplitude in section 4.3 also apply for the output of an SQF. The DOCP is an odd filter with isotropic bandpass characteristic, and therefore realizes a filter that has been thought to be impossible [82]. In figure 4.16 the SQF is visualized in the spatial domain and in the frequency domain.

The spherical quadrature filters are *band-selective* filters, which means that they extend the split of identity by a decomposition into partial signals (see section 3.2 and figure 3.3). Since the SQF fulfill the scaling equation, their contribution to the signal energy is independent of the passband number. This leads to the following conclusion: just as the Riesz transform is dominated by the main i1D signal, the signal is dominated by the partial signal with the highest energy. This partial signal is represented by a bandpass of appropriate scale. Obviously, this aspect leads to the discussion about the *intrinsic scale* of an image.

[17] Note that the Riesz transform yields two components, so that the original function plus its Riesz transform form a triple.

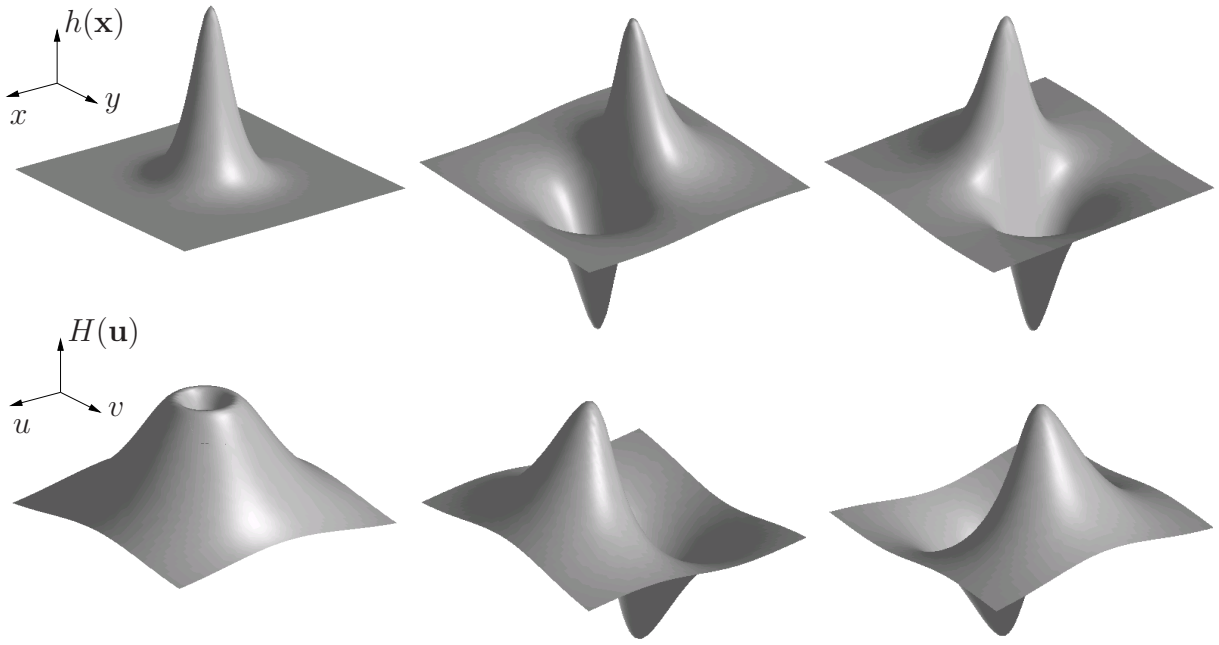

Fig. 4.16: SQF in the spatial domain (top row) and in the frequency domain (bottom row). Note that the impulse response in the upper left has a zero DC component which is hard to see in this figure.

4.6.3 The Intrinsic Scale

Basically, *the* intrinsic scale of an image does not exist because the dominant local frequency in an image varies with its spatial position. Although this effect could be compensated by introducing an approach for a *local intrinsic scale*, this would not be satisfying either. An image is a superposition of many partial signals with different scales. Even at one certain position several relevant scales may occur. Hence, the intrinsic-scale-map of an image is a multi-valued mapping which is not constant in general. One should rather speak of a *set of local intrinsic scales* instead.

The set of intrinsic scales at a certain position is expected to have only few elements. Even in very detailed images only a small number of bandpass *channels*[18] are relevant and in many cases there is only one dominating channel (in many cases it is the DC channel). Of course this hypothesis cannot be proofed, but it can be verified by experiments. One example is shown in figure 4.17. In this figure the graph of the local amplitude is drawn against the scale for four different points. The point number 3 (chin) does not show any significant maximum in the amplitude. This corresponds to a partial signal which is locally dominated by its DC component.

[18] The channel representation [60, 53] is a generalization of quantization [52]. The basic idea is to have intersecting support for neighbored features.

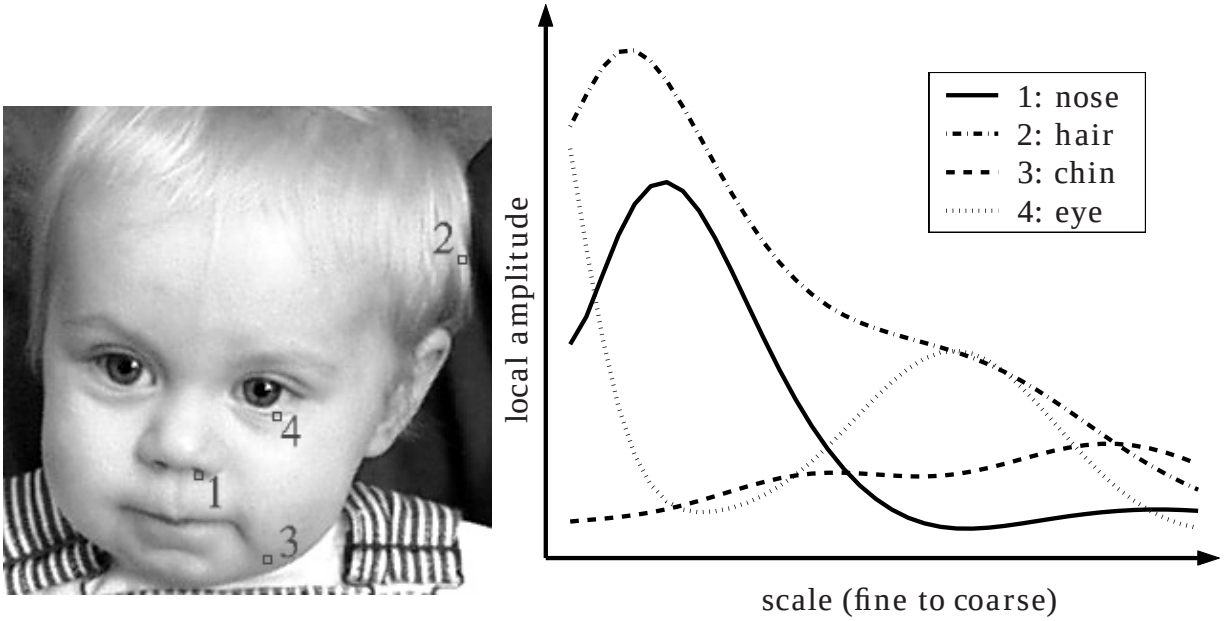

Fig. 4.17: Distribution of the local energy in scale-space at four points of an image.

The point number 1 (nose) shows one sharp maximum. This corresponds to one intrinsic scale. The point number 4 however, shows two maxima. Hence, there are two intrinsic scales at that position. The point number 2 (hair) also shows one sole local maximum, but there is a second 'nearly maximum' which indicates that in the neighborhood the energy for that band might be high. This observation yields another consequence: the local intrinsic scale sets must be smooth to some extent. Further properties of the intrinsic scale map and a scale-detection algorithm based on SQFs are discussed in chapter 6.

In the subsequent chapter, the approach of the Poisson scale-space and SQF is extended for anisotropic settings. The need for introducing anisotropies is quite obvious when considering i1D structures. Whereas the phase approach of the monogenic signal and the SQF is suited for representing i1D structures, the isotropy of the filters yields a suboptimal estimation of local properties of i1D signals. The filter extent parallel to the i1D structure can be enlarged without changing the result of the estimation. On the other hand, such a filter is more stable against noise. A second reason for introducing anisotropies is the extension of the monogenic signal for describing i2D structures, which is not possible without having anisotropic filters.

4.7 Summary of Chapter 4

In this chapter the 1D analytic signal has been derived by means of 2D harmonic vector fields. The second coordinate has been identified as a scale-parameter. The analytic signal on each scale is obtained by convolution with the 1D Poisson filter and the conjugate 1D Poisson filter. The derivation directly generalizes to 3D vector fields which yields the Riesz transform as the 2D generalization of the Hilbert transform. The combination of a signal with its Riesz transform is a 2D generalization of the analytic signal which is called the monogenic signal. Several properties of the derived approach have been investigated, especially the estimation of the local phase. The phase interpretation is established by a theorem that relates the Riesz transform and the Hilbert transform by means of the Radon transform.

The third coordinate which is involved in the derivation of the Riesz transform can be identified with a scale parameter. The corresponding linear scale-space which is obtained by the 2D Poisson filter has been investigated with respect to some scale-space axiomatics and it has been compared to the Gaussian scale-space. The bandpass filters obtained as the difference of Poisson filters establish a continuous wavelet transform and they have been combined with the corresponding difference of conjugate Poisson filters to form the spherical quadrature filters. By means of these filters, the notion of intrinsic scale has been discussed.

The most important results of this chapter are summarized in figure 4.2, figure 4.4, table 4.1, figure 4.8, and table 4.2.

filter	formula
DOP (impulse resp.)	$b_{s_0,\lambda,k}(\mathbf{x}) = \frac{s_0\lambda^k}{2\pi(\mathbf{x}^2+(s_0\lambda^k)^2)^{3/2}} - \frac{s_0\lambda^{k-1}}{2\pi(\mathbf{x}^2+(s_0\lambda^{k-1})^2)^{3/2}}$
DOP (frequency resp.)	$B_{s_0,\lambda,k}(\mathbf{u}) = \exp(-2\pi\lvert\mathbf{u}\rvert s_0\lambda^k) - \exp(-2\pi\lvert\mathbf{u}\rvert s_0\lambda^{k-1})$
DOCP (impulse resp.)	$c_{s_0,\lambda,k}(\mathbf{x}) = \frac{\mathbf{x}\mathbf{e}_3}{2\pi(\mathbf{x}^2+(s_0\lambda^k)^2)^{3/2}} - \frac{\mathbf{x}\mathbf{e}_3}{2\pi(\mathbf{x}^2+(s_0\lambda^{k-1})^2)^{3/2}}$
DOCP (frequency funct.)	$C_{s_0,\lambda,k}(\mathbf{u}) = \frac{\mathbf{u}}{\lvert\mathbf{u}\rvert} I_2^{-1}(\exp(-2\pi\lvert\mathbf{u}\rvert s_0\lambda^k) - \exp(-2\pi\lvert\mathbf{u}\rvert s_0\lambda^{k-1}))$

Tab. 4.2: Overview of formulas for the spherical quadrature filter.

Chapter 5

ADAPTION OF 2D QUADRATURE FILTERS TO THE LOCAL INTRINSIC DIMENSION

> '*The mathematical sciences particularly exhibit order, symmetry, and limitation; and these are the greatest forms of the beautiful.*'
>
> Aristotle (384-322 b.C.)

In chapter 4 an isotropic approach for a 2D generalized analytic signal is introduced. The local phase definition is based on the underlying *local intrinsically 1D signal*, where the latter is obtained according to an L^2-approximation of the original local signal (see discussion about the slice theorem in section 4.4). The isotropic approach yields two drawbacks. On the one hand, the monogenic phase is not suited for encoding locally i2D structures. On the other hand, the local properties of a signal which *is* i1D are not estimated in an optimal manner, since the smoothing of the applied bandpass does not take into account the anisotropy of the filter input. The estimation will be improved if the bandpass smoothes more strongly parallel to the i1D structure and smoothes less orthogonal to the structure (see e.g. [28]). Hence, the monogenic phase of the isotropic SQF is well-suited for i1D signals whereas the amplitude response of the SQF is optimal for isotropic signals, i.e., i2D or i0D signals.

An optimal i1D filter should be composed of the phase approach introduced in the previous chapter and an *anisotropic bandpass*. Conversely, an optimal filter for the remaining signals (i0D and i2D) should consist of a new *extended phase approach* and an isotropic bandpass (as introduced in chapter 4). The two kinds of locally non-i1D signals can be processed by the same method since in the case of local i2D signals, the relevant spectral components include high frequency components for every orientation, whereas in the i0D case only low frequency components exist. Hence, the difference between locally i0D signals and locally i2D signals can be reduced to the radial spectral distribution, or, in other words, to the local scale (see end of section 4.5).

The first three sections of this chapter are about optimized i1D filters and deal with the question, how the angular spectral distribution of i1D signals can be taken into account in order to improve the estimation of their local properties. In the remaining part of this chapter, an extension of the 2D phase approach for i2D signals is developed, which allows to analyze the 2D structure by means of a symmetry decomposition. The new method can be considered as an i2D 2D analytic signal although such an approach cannot be *complete* unless it consists of an infinite number of filters (see section 1.1).

5.1 Introducing Local Anisotropies

Starting point for extending the spherical quadrature filters to an anisotropic support is to introduce a *local metric* on the underlying space. In the current context, a local metric is such a one which varies globally in the embedding 3D space, but for a local neighborhood it is constant.

5.1.1 The Metric Tensor

The classical method for describing the metric of a vector space is to introduce a *metric tensor* $\mathbf{G}$, such that the scalar product of a vector $\mathbf{v} = (v_1, v_2, v_3)^{\mathrm{T}} \in \mathbb{R}^3$ with itself is replaced by a symmetric bilinear form according to

$$\langle \mathbf{v}, \mathbf{v} \rangle_{\mathbf{G}} = \mathbf{v}^{\mathrm{T}} \mathbf{G} \mathbf{v} = (v_1\, v_2\, v_3) \begin{bmatrix} g_{11} & g_{12} & g_{13} \\ g_{12} & g_{22} & g_{23} \\ g_{13} & g_{23} & g_{33} \end{bmatrix} \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} \quad , \tag{5.1}$$

where $\mathbf{G}$ is positive definite and symmetric. Hence, the introduction of a metric tensor is equivalent to the definition of a more general symmetric, positive definite bilinear form on the vector space. Geometric algebras and harmonic calculus can be defined for those vector spaces as well [58]. The question is, however, how to represent these generalized approaches in Euclidean space again. This representation is necessary since the image data is always given in Euclidean 2D coordinates and also the 3rd coordinate (the scale vs. the intensity axis) is in fixed relation to the image coordinates. The latter fact can be formulated more concretely in the following way. Obviously, it is not reasonable to mix image coordinates with the intensity coordinate [81]. Accordingly, the components g_{13} and g_{23} are zero. The only non-zero component involving the third coordinate is g_{33} which is similar to a *homogeneous coordinate*. Since the intensity coordinate should remain in the original scale (otherwise artificial enhancements of contrast would be introduced), the tensor is divided by g_{33}.

The metric tensor will be estimated from the image itself. Since the grayscale is attached to the third coordinate (see figure 3.4), the local grayscale dynamics is reflected in the component g_{33}. The metric tensor is obtained such that it normalizes distances in the embedded signal, and therefore it compensates the local grayscale dynamics. Hence, locally high grayscale dynamics implies a low value of g_{33} and vice versa. Since the scale coordinate is also identified with the third coordinate, a low value for g_{33} implies a slower change in scale, i.e., the scale remains fine, and a high value for g_{33} yields a fast change in scale, i.e., the scale becomes coarse very quickly. Hence, if the image is of locally low contrast (or nearly i0D), the scale is chosen to be coarser whereas for high contrast the scale is chosen to be finer. In the deformed space however, a low value for g_{33} results in *damping* the local intensities whereas a high value for g_{33} *amplifies* the local intensities. Hence, a *local gain control* is introduced, i.e., a weighting by the local dynamics (see e.g. [52]). Hence, low local contrast implies fast changes in scale in the original space and an amplification of intensities in the deformed space. High local contrast implies slow changes in scale in the original space and a damping of intensities in the deformed space. The change of scale in the deformed space is constant.

Considering the component g_{33} in the light of a homogeneous coordinate reduces the whole problem to the 2D tensor $\mathbf{G}' = g_{33}^{-1}\mathbf{G}$. This tensor can be estimated by the structure tensor.[1] The latter contains three kinds of information:

1. local isotropic scale (amplitude)
2. local isotropy (coherence)
3. local orientation

These three different features are obtained by an eigenvalue decomposition of the tensor and they can be visualized as a geometric object: an ellipse (see figure 5.1). The local orientation vector $\mathbf{n}$ is obtained as the first eigenvector of the tensor and corresponds to the larger eigenvalue. Any vector parallel to $\mathbf{n}$ is multiplied by λ_1, the square root of the first eigenvalue. Hence the corresponding energy (the square of the vector) is multiplied by λ_1^2. The orthogonal vector $\mathbf{n}I_2^{-1}$ spans the space of vectors which are multiplied by λ_2, the square root of the second eigenvalue. The isotropic scale is given by the square root of the tensor determinant (see next section) which is identical to the area of the ellipse up to the constant factor π.

According to [71] geometric algebra allows coordinate-free computations. Indeed, the computations using the metric tensor can be simplified by transferring the problem to geometric algebra. The idea is the following: all vectors in the 3D vector space

[1] The metric tensor $\mathbf{G}'$ differs from the induced metric tensor according to [134]. The latter is identical to the sum of the structure tensor according to (3.48) and the identity tensor.

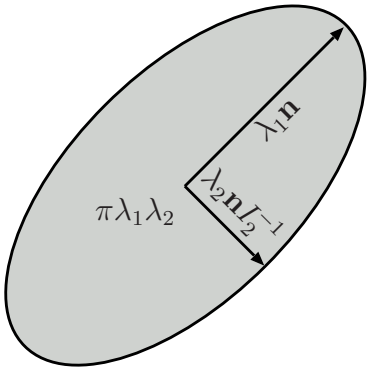

Fig. 5.1: An ellipse representing the structure tensor and the metric tensor. The vector $\mathbf{n}$ indicates the main orientation, the two principal axes correspond to the maximal and minimal energy. The area of the ellipse is proportional to the isotropic scale since $\lambda_1^2\lambda_2^2$ is the determinant of the tensor $\mathbf{G}'$.

are projected onto the normalized principal axes of the ellipsoid, i.e.,

$$\mathbf{v}_j = (\mathbf{n}_j \cdot \mathbf{v})\mathbf{n}_j \qquad \text{for } j = 1, \ldots, 3 \ . \tag{5.2}$$

The bilinear form in the space with metric $\mathbf{G}$ is then simplified to be

$$\langle \mathbf{v}, \mathbf{v} \rangle_{\mathbf{G}} = \sum_{j=1}^{3} (\lambda_j' \mathbf{v}_j)^2 \ , \tag{5.3}$$

where λ_j' are the lengths of the principal axes. In the case of images, the third principal axis is fixed to be $\lambda_3'\mathbf{e}_3$, and hence all (normalized) axes are given by just one vector $\mathbf{n}$ in the I_2-plane: $\mathbf{n}_3 = \mathbf{e}_3$, $\mathbf{n}_2 = \mathbf{n}I_2^{-1}$, and $\mathbf{n}_1 = \mathbf{n}$. Thinking of $\mathbf{e}_3$ being a homogeneous coordinate, the three axes-lengths of the ellipsoid are reduced to two: $\lambda_1 = \lambda_1'/\lambda_3'$ and $\lambda_2 = \lambda_2'/\lambda_3'$. Hence, the rather abstract metric tensor is represented as an ellipse (see again figure 5.1).

5.1.2 The Laplace Equation in the Deformed Space

The question arises whether it is reasonable to consider the 3D space as a homogeneous 2D space. In the context of the Laplace equation (2.45), it is actually reasonable since this differential equation is a homogeneous PDE and therefore independent of an absolute scaling. Applying the Laplace equation in the deformed space yields the *anisotropic Laplace equation*

$$\Delta_{\mathbf{G}} p = \langle \nabla_3, \nabla_3 \rangle_{\mathbf{G}} p = (\lambda_1^2 (\mathbf{n} \cdot \nabla_3)^2 + \lambda_2^2 ((-\mathbf{n}I_2) \cdot \nabla_3)^2 + (\mathbf{e}_3 \cdot \nabla_3)^2) p = 0 \tag{5.4}$$

in the Euclidean space. Separating the 3D gradient operator into the 2D gradient operator and the second s-derivative operator, this equation reads

$$(\lambda_1^2(\mathbf{n}\cdot\nabla_2)^2 - \lambda_2^2(\mathbf{n}\wedge\nabla_2)^2)p = -\frac{\partial^2}{\partial s^2}p \ , \tag{5.5}$$

where the projection $((-\mathbf{n}I_2)\cdot\nabla_2)(-\mathbf{n}I_2)$ is rewritten as a rejection

$$\begin{aligned}((-\mathbf{n}I_2)\cdot\nabla_2)(-\mathbf{n}I_2) &= \frac{1}{2}(\mathbf{n}I_2\nabla_2 + \nabla_2\mathbf{n}I_2)(\mathbf{n}I_2)\\ &= \frac{1}{2}(\mathbf{n}I_2\nabla_2\mathbf{n}I_2 + \nabla_2\mathbf{n}I_2\mathbf{n}I_2)\\ &= \frac{1}{2}(-\mathbf{n}\nabla_2\mathbf{n} + \nabla_2\mathbf{n}\mathbf{n})\\ &= \frac{1}{2}(\nabla_2\mathbf{n} - \mathbf{n}\nabla_2)\mathbf{n} = (\nabla_2\wedge\mathbf{n})\mathbf{n} = \mathbf{n}(\mathbf{n}\wedge\nabla_2) \ .\end{aligned} \tag{5.6}$$

The fundamental solution of (5.5) is given as

$$p_3(\mathbf{x},s) = -\frac{1}{2\pi\sqrt{\lambda_2^2(\mathbf{n}\cdot\mathbf{x})^2 - \lambda_1^2(\mathbf{n}\wedge\mathbf{x})^2 + \lambda_1^2\lambda_2^2 s^2}} \ . \tag{5.7}$$

The proof is quite short in geometric algebra. It is based on the two-fold application of the *anisotropic gradient operator*

$$\nabla_{\mathbf{G}} = \lambda_1\mathbf{n}(\mathbf{n}\cdot\nabla_2) + \lambda_2\mathbf{n}(\mathbf{n}\wedge\nabla_2) + \mathbf{e}_3\frac{\partial}{\partial s} \ . \tag{5.8}$$

Knowing that $\mathbf{n}\cdot\nabla_2$ is just the *directed derivative* along $\mathbf{n}$ and that $\mathbf{n}\wedge\nabla_2\, I_2$ is the directed derivative along the orthogonal direction $\mathbf{n}I_2^{-1}$, most of the mixed derivatives are obviously zero. In order to have a shorter notation, the following abbreviation is used[2] $B(\mathbf{x},s) = \lambda_2^2(\mathbf{n}\cdot\mathbf{x})^2 - \lambda_1^2(\mathbf{n}\wedge\mathbf{x})^2 + \lambda_1^2\lambda_2^2 s^2$. Applying the anisotropic gradient operator to p_3 once yields

$$\begin{aligned}\frac{\partial}{\partial_s}p_3(\mathbf{x},s) &= \lambda_1\lambda_2\frac{\lambda_1\lambda_2 s}{2\pi\sqrt{B(\mathbf{x},s)}^3}\\ \lambda_1\mathbf{n}(\mathbf{n}\cdot\nabla_2)\,p_3(\mathbf{x},s) &= \lambda_1\lambda_2\frac{\lambda_2\mathbf{n}(\mathbf{n}\cdot\mathbf{x})}{2\pi\sqrt{B(\mathbf{x},s)}^3}\\ \lambda_2\mathbf{n}(\mathbf{n}\wedge\nabla_2)\,p_3(\mathbf{x},s) &= \lambda_1\lambda_2\frac{\lambda_1\mathbf{n}(\mathbf{n}\wedge\mathbf{x})}{2\pi\sqrt{B(\mathbf{x},s)}^3}\end{aligned}$$

[2] Actually, B is just the bilinear form induced by the inverse metric tensor.

and the second application gives

$$\begin{aligned}
\frac{\partial^2}{\partial_s^2} p_3(\mathbf{x}, s) &= \lambda_1^2\lambda_2^2 \left(\frac{1}{2\pi\sqrt{B(\mathbf{x}, s)}^3} - \frac{3\lambda_1^2\lambda_2^2 s^2}{2\pi\sqrt{B(\mathbf{x}, s)}^5} \right) \\
\lambda_1^2(\mathbf{n}\cdot\nabla_2)^2 p_3(\mathbf{x}, s) &= \lambda_1^2\lambda_2^2 \left(\frac{1}{2\pi\sqrt{B(\mathbf{x}, s)}^3} - \frac{3\lambda_2^2(\mathbf{n}\cdot\mathbf{x})^2}{2\pi\sqrt{B(\mathbf{x}, s)}^5} \right) \\
-\lambda_2^2(\mathbf{n}\wedge\nabla_2)^2 p_3(\mathbf{x}, s) &= \lambda_1^2\lambda_2^2 \left(\frac{1}{2\pi\sqrt{B(\mathbf{x}, s)}^3} + \frac{3\lambda_1^2(\mathbf{n}\wedge\mathbf{x})^2}{2\pi\sqrt{B(\mathbf{x}, s)}^5} \right)
\end{aligned}$$

so that the sum is zero:

$$\begin{aligned}
\nabla^2_{\mathbf{G}}\, p_3(\mathbf{x}, s) &= \lambda_1^2\lambda_2^2 \left(\frac{3}{2\pi\sqrt{B(\mathbf{x}, s)}^3} - \frac{3(\lambda_2^2(\mathbf{n}\cdot\mathbf{x})^2 - \lambda_1^2(\mathbf{n}\wedge\mathbf{x})^2 + \lambda_1^2\lambda_2^2 s^2)}{2\pi\sqrt{B(\mathbf{x}, s)}^5} \right) \\
&= \lambda_1^2\lambda_2^2 \left(\frac{3}{2\pi\sqrt{B(\mathbf{x}, s)}^3} - \frac{3B(\mathbf{x}, s)}{2\pi\sqrt{B(\mathbf{x}, s)}^5} \right) = 0 \quad .
\end{aligned}$$

Hence, (5.7) is the fundamental solution of the anisotropic Laplace equation (5.5). Applying the anisotropic gradient operator to p_3 yields the vector field consisting of the *anisotropic Poisson kernel* and its conjugate

$$\mathbf{f}_3(\mathbf{x}, s) = \nabla_{\mathbf{G}} p_3(\mathbf{x}, s) = \lambda_1\lambda_2 \frac{\lambda_2\mathbf{n}(\mathbf{n}\cdot\mathbf{x}) + \lambda_1\mathbf{n}(\mathbf{n}\wedge\mathbf{x}) + \lambda_1\lambda_2 s}{2\pi\sqrt{B(\mathbf{x}, s)}^3} \quad . \tag{5.9}$$

5.1.3 Interpretation in the Fourier Domain

In order to get an appropriate interpretation of the previous results, the Fourier domain representation of the induced vector field is considered in the following. Applying the affine theorem for anisotropic dilations (3.38), the frequency responses of the anisotropic Poisson kernel and its conjugate are obtained, using the abbreviation $|\mathbf{u}'| = \sqrt{\lambda_2^{-2}(\mathbf{n}\cdot\mathbf{u})^2 - \lambda_1^{-2}(\mathbf{n}\wedge\mathbf{u})^2}$, as

$$\mathcal{F}_2\{\mathbf{f}_3 \cdot \mathbf{e}_3\}(\mathbf{u}, s) = \exp\left(-2\pi|\mathbf{u}'|\lambda_1\lambda_2 s\right) \tag{5.10}$$

$$\mathcal{F}_2\{\mathbf{f}_3 \wedge \mathbf{e}_3\}(\mathbf{u}, s) = H_2(\mathbf{u}) \exp\left(-2\pi|\mathbf{u}'|\lambda_1\lambda_2 s\right) \tag{5.11}$$

where $H_2(\mathbf{u})$ is now the *anisotropic Riesz transform*:

$$H_2(\mathbf{u}) = \frac{\lambda_2^{-1}\mathbf{n}(\mathbf{n}\cdot\mathbf{u}) + \lambda_1^{-1}\mathbf{n}(\mathbf{n}\wedge\mathbf{u})}{|\mathbf{u}'|} I_2^{-1} \quad . \tag{5.12}$$

Hence, the calculations on the previous two pages result in frequency responses that are quite simple to interpret. The introduction of the deformed space just yields a

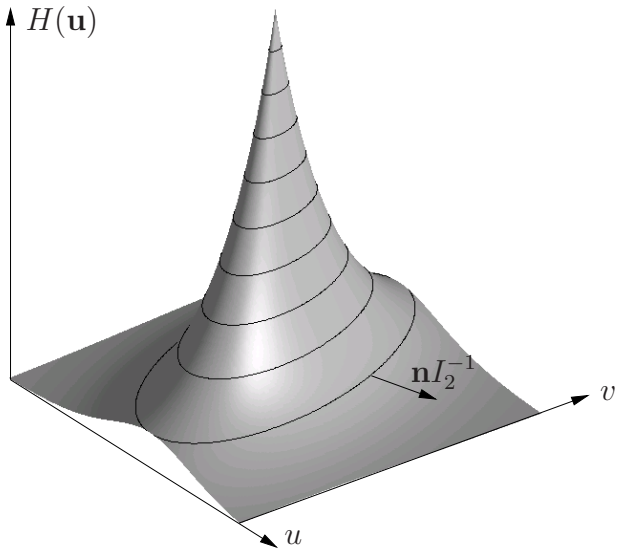

Fig. 5.2: Frequency response of the anisotropic Poisson kernel. The lines of constant amplitude are ellipses. The area of an isoplethal ellipse indicates the total smoothing whereas the aspect-ratio indicates the isotropy (here: $\lambda_1 = \sqrt{2}$ and $\lambda_2 = \sqrt{1/2}$). The orientation of the ellipse is given by the unit vector $\mathbf{n}$.

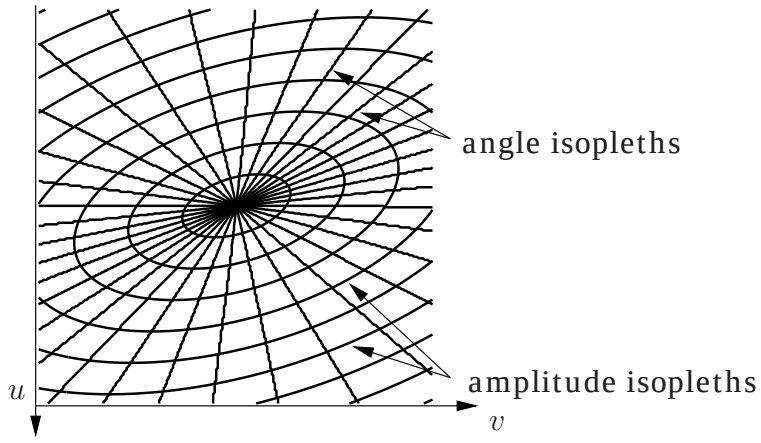

Fig. 5.3: Representation of the anisotropic Riesz transform as a bunch graph. The rotation field is adapted to the isopleths in figure 5.2 which are visualized by the ellipses. Instead of representing the Riesz transform by a vector field as in figure 4.5, the radial rays correspond to the isopleths of the angle of the anisotropic Riesz transform.

more general smoothing kernel with ellipses as iso-lines (see figure 5.2) and a Riesz transform that evaluates the deformed angle (see figure 5.3). The interpretation of the Riesz transform for $\lambda_1 = \lambda_2 = 1$ is exactly the one of the isotropic case (see chapter 4). For $\lambda_2 = \varepsilon = \lambda_1^{-1}$ and $\varepsilon \to 0$ however, the Riesz transform converges to the partial Hilbert transform! This can be simply verified by considering the limit:

$$\begin{aligned}\lim_{\varepsilon\to 0} H_2(\mathbf{u}) &= \lim_{\varepsilon\to 0} \frac{\varepsilon^{-1}\mathbf{n}(\mathbf{n}\cdot\mathbf{u}) + \varepsilon\mathbf{n}(\mathbf{n}\wedge\mathbf{u})}{\sqrt{\varepsilon^{-2}(\mathbf{n}\cdot\mathbf{u})^2 - \varepsilon^2(\mathbf{n}\wedge\mathbf{u})^2}} I_2^{-1} \\ &= \lim_{\varepsilon\to 0} \frac{\varepsilon^{-1}\mathbf{n}(\mathbf{n}\cdot\mathbf{u})}{\sqrt{\varepsilon^{-2}(\mathbf{n}\cdot\mathbf{u})^2}} I_2^{-1} \\ &= \mathbf{n}\frac{\mathbf{n}\cdot\mathbf{u}}{|\mathbf{n}\cdot\mathbf{u}|} I_2^{-1} = \mathbf{n} I_2^{-1}\,\mathrm{sign}(\mathbf{n}\cdot\mathbf{u}) \quad .\end{aligned} \tag{5.13}$$

The impact of this result to the theory of quadrature filter design is fundamental: *The partial Hilbert transform and the Riesz transform are the complementary cases in the design of general anisotropic quadrature filters.* The term *complementary* in the statement above refers to the intrinsic dimension of the considered signals. Whereas the Riesz transform is appropriate for 'ideal' i2D signals, i.e., signal which have a constant angular energy distribution, the Hilbert transform is optimal for 'ideal' i1D signals, i.e., signals which are zero everywhere except for their main orientation.

The statement about the complementarity of the Hilbert transform and the Riesz transform leads to several conclusions. First, it is not reasonable to combine the partial Hilbert transform with an isotropic bandpass within the framework of the Laplace equation. The reason for this is that the Hilbert transform is obtained by taking the limit $\varepsilon \to 0$ which implies also to take the limit of the lowpass (5.10). The latter limit never results in an isotropic lowpass. Assuming that $\lambda_1\lambda_2$ is a finite constant, the lowpass (5.10) degenerates to a line-impulse orthogonal to $\mathbf{n}$. If λ_2 tends to zero and λ_1 is a constant greater than zero, $\lambda_1\lambda_2$ also tends to zero, which corresponds to an infinite fine isotropic scale. In this case the limit of the lowpass (5.10) is a function of $\mathbf{n}\cdot\mathbf{u}$ solely, corresponding to the Hilbert transform (5.13). Hence, using the partial Hilbert transform implies to use a bandpass which only depends on the frequency component parallel to the main orientation. Using other amplitude responses instead, as it is often done in practice, yields amplitude and phase distortions, see also section 1.1.

A second consequence is that the Radon transform can be considered as the degenerated case of an anisotropic Poisson kernel. The Radon transform in section 4.4 is just the output of the degenerated Poisson kernel applied for all possible orientations. Practical applications of this knowledge (e.g. using a non-degenerated Poisson kernel with a high degree of anisotropy instead of the Radon transform) might be possible, but they are out of the scope of this thesis.

5.2 Approximation of the Local Metric

Assuming that the local metric tensor $\mathbf{G}$ is known, the anisotropic Poisson kernel and its conjugate can be used to design *elliptic quadrature filters* (EQF) by taking the difference of anisotropic Poisson kernels and conjugate Poisson kernels at different scales. Unfortunately, the filter is not shift invariant since the metric depends on the spatial position. As a consequence, the computational load for computing the EQF is raising from $O(N)$ to $O(N^2)$ (N being the 2D signal size). To reduce the complexity, the following method is used in this section. Instead of computing the exact EQF response, a general approximation of the EQF filter by means of a Taylor series expansion is derived. The resulting filter scheme reduces the complexity for computing the EQF output to $O(N^{3/2})$ since the applied basis filters are independent of the local isotropy, they only depend on the local orientation.

5.2.1 Taylor Expansion of the Poisson Kernel

In order to reduce the computational complexity, the anisotropic Poisson kernel (5.10) is approximated by a first order Taylor expansion. In order to keep the derivation simple, the following abbreviations are used in the sequel.

$$c = \frac{\lambda_2}{\lambda_1} \tag{5.14}$$

$$\mathbf{u}' = c^{-1/2}\,\mathbf{n}(\mathbf{n}\cdot\mathbf{u}) + c^{1/2}\,\mathbf{n}(\mathbf{n}\wedge\mathbf{u}) \quad \text{so that} \quad |\mathbf{u}'| = \sqrt{c^{-1}(\mathbf{n}\cdot\mathbf{u})^2 - c(\mathbf{n}\wedge\mathbf{u})^2} \tag{5.15}$$

$$s' = \sqrt{\lambda_1\lambda_2}s \tag{5.16}$$

$$L(c) = \exp(-2\pi|\mathbf{u}'|s') \qquad \text{according to (5.10)} \tag{5.17}$$

$$H(c) = \frac{\mathbf{u}'}{|\mathbf{u}'|}I_2^{-1} \qquad \text{according to (5.12)} \tag{5.18}$$

The scale parameter s' from (5.16) is the geometric mean of the maximal and the minimal scale parameter. It can be considered as the *isotropic scale parameter*, or *total scale parameter*, which yields a circle in the image plane that has the same area as the ellipse which is created by $\lambda_1 s$ and $\lambda_2 s$ (see also figure 5.1). The ratio of the principal axes, c, is called the *local isotropy*. Note that the arguments $\mathbf{x}$ and $\mathbf{u}$ are omitted for the subsequent derivations.

The c-derivative of $L(c)$ reads

$$\begin{aligned} L_c(c) = \frac{\partial L(c)}{\partial c} &= L(c)\frac{\partial}{\partial c}(-2\pi|\mathbf{u}'|s') \\ &= \frac{\pi s'(c^{-2}(\mathbf{n}\cdot\mathbf{u})^2 + (\mathbf{n}\wedge\mathbf{u})^2)}{|\mathbf{u}'|}L(c) \end{aligned}$$

and if α denotes the angle between $\mathbf{n}$ and $\mathbf{u}$

$$L_c(c) = \frac{\pi s' |\mathbf{u}|^2 (c^{-2}\cos^2\alpha - \sin^2\alpha)}{|\mathbf{u}'|} L(c) \ .$$

For $c = 1$ this simplifies to (note that $\mathbf{u}' = \mathbf{u}$ in this case):

$$L_c(1) = \pi s' |\mathbf{u}| \cos(2\alpha) L(1) \ . \tag{5.19}$$

Hence, the first order Taylor expansion of $L(c)$ in the point $c = 1$ is given by

$$L(c) \approx L(1) + L_c(1)(c-1) = (1 + \pi s' |\mathbf{u}| (c-1)\cos(2\alpha)) \exp(-2\pi |\mathbf{u}| s') \ . \tag{5.20}$$

By means of this expansion, the anisotropic Poisson kernel can be approximated by one isotropic term and one anisotropic term. This reduces the computational load by one order since instead of computing a different filter output for each different isotropy factor c, only a simple linear combination of two basis filters must be evaluated. Furthermore, the anisotropic term, i.e., the $\cos(2\alpha)$-term, can easily be calculated by a steered filter (see next section).

The accuracy of the approximation depends on the isotropy parameter c. If c is close to one, the approximation is quite good, see figure 5.4. To be more concrete, the

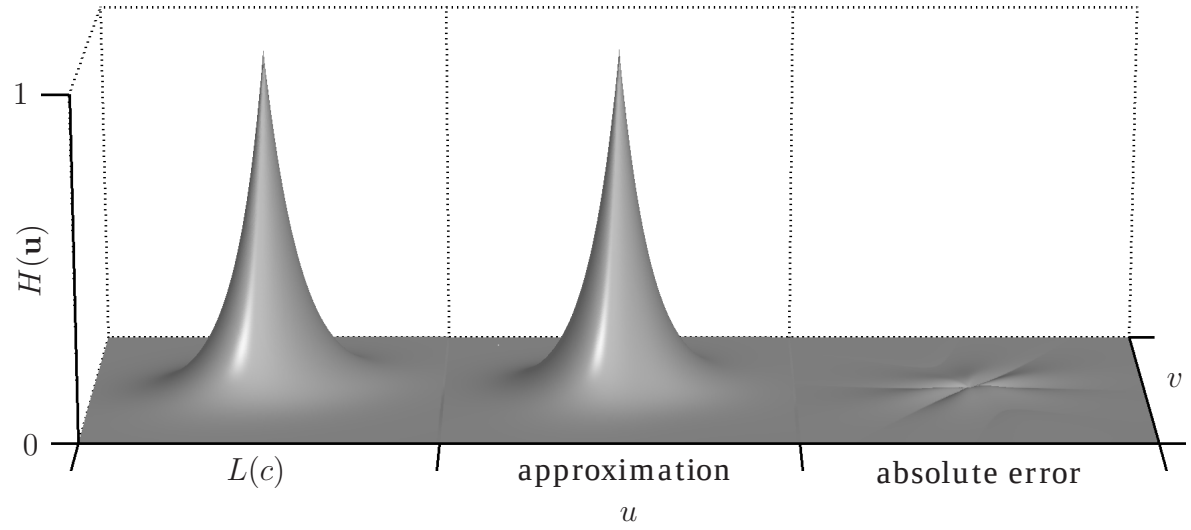

Fig. 5.4: Approximation of the anisotropic Poisson kernel by a first-order Taylor expansion. Left: original frequency response for $c = 2/3$, center: approximation, right: absolute error of the approximation.

relative energy of the approximation error is independent of the scale parameter s' and it is less than 0.2% of the signal energy for $c = 2/3$. For $c = 1/2$ the error is less than 1.3% of the signal energy which is still an acceptable error for many

applications. A ratio of $c = 1/2$ corresponds to a coherence of $((2-1/2)/(2+1/2))^2 = \frac{9}{25}$ where the range of the coherence is in $[0, 1]$ (for the relation to the coherence of the structure tensor see below).

5.2.2 Taylor Expansion of the Riesz Transform

In order to approximate the anisotropic conjugate Poisson kernel, the anisotropic Riesz transform has to be expressed by a Taylor series expansion. The c-derivative of $H(c)$ is obtained as

$$\begin{aligned} H_c(c) = \frac{\partial H(c)}{\partial c} &= \frac{\frac{\partial \mathbf{u}'}{\partial c}|\mathbf{u}'| - \mathbf{u}'\frac{\partial |\mathbf{u}'|}{\partial c}}{|\mathbf{u}'|^2} I_2^{-1} \\ &= \left(\frac{\mathbf{n}(-c^{-3/2}\,\mathbf{n}\cdot\mathbf{u} + c^{-1/2}\,\mathbf{n}\wedge\mathbf{u})}{2|\mathbf{u}'|} + \frac{\mathbf{u}'(c^{-2}(\mathbf{n}\cdot\mathbf{u})^2 + (\mathbf{n}\wedge\mathbf{u})^2)}{2|\mathbf{u}'|^3} \right) I_2^{-1} \\ &= \frac{1}{2}\left(\mathbf{n}(-c^{-3/2}\cos\alpha + c^{-1/2}\sin\alpha I_2) + \frac{\mathbf{u}'}{|\mathbf{u}'|}(c^{-2}\cos^2\alpha - \sin^2\alpha) \right) I_2^{-1} \quad . \end{aligned}$$

For $c = 1$ this equation reduces to

$$H_c(1) = \frac{1}{2}\left(-\mathbf{n}\exp(-I_2\alpha) + \frac{\mathbf{u}}{|\mathbf{u}|}\cos(2\alpha) \right) I_2^{-1} \quad . \tag{5.21}$$

The term $\mathbf{n}\exp(-I_2\alpha)$ is equal to $\frac{\mathbf{nun}}{|\mathbf{u}|}$, and hence the first order Taylor expansion of $H(c)$ in the point $c = 1$ is given by

$$H(c) \approx H(1) + H_c(1)(c-1) = \frac{2\mathbf{u} - (\mathbf{nun} - \mathbf{u}\cos(2\alpha))(c-1)}{2|\mathbf{u}|} I_2^{-1} \quad . \tag{5.22}$$

The approximation of the Riesz transform by this expansion is illustrated in figure 5.5.

Using this expansion, the conjugate Poisson kernel can be approximated by the following first order expansion:

$$H(c)L(c) \approx H(1)L(1) + (H_c(1)L(1) + H(1)L_c(1))(c-1) \quad . \tag{5.23}$$

In figure 5.6 the two components of the frequency response of the conjugate Poisson kernel, their approximations, and their error functions are illustrated. In the case of the conjugate Poisson kernel, the error of the approximation is doubled with respect to the Poisson kernel approximation since the errors of both Taylor expansions accumulate. In the case $c = 2/3$ the error is less than 0.4% and in the case $c = 1/2$ the error is less than 2.6%. Hence, using the Taylor expansions of the Poisson kernel and its conjugate kernel for two different scales s', the elliptic quadrature

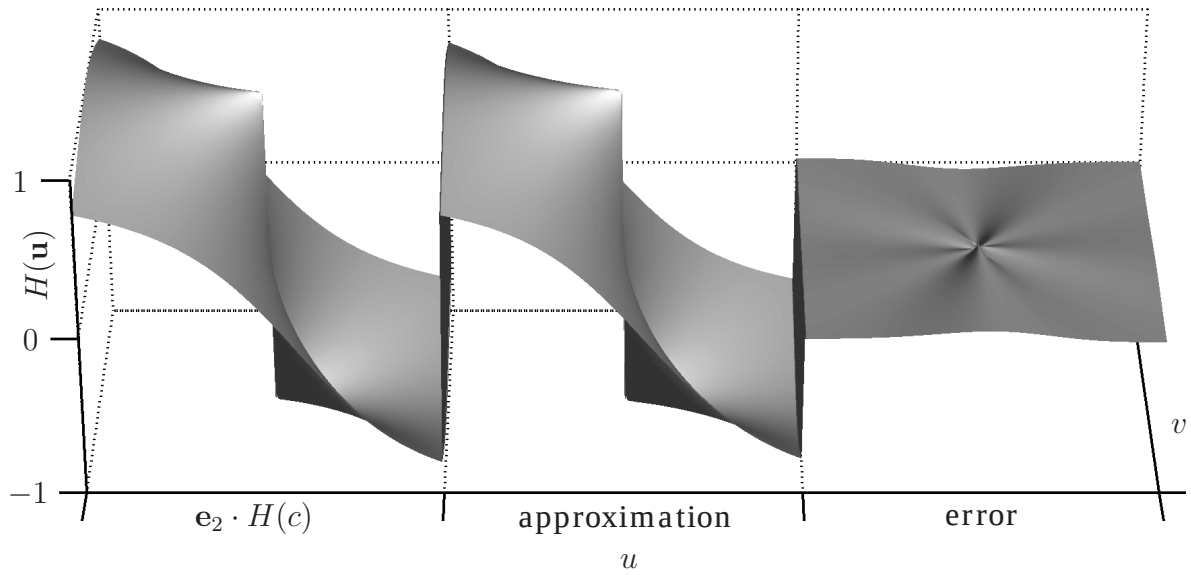

Fig. 5.5: Approximation of the anisotropic Riesz transform by expansion (5.22). Left: original frequency response of one component of $H(c)$ for $c = 2/3$, center: approximation, right: error of the approximation.

filter which has been introduced in section 5.1 can be applied with a computational complexity of $O(N^{3/2})$. A further decrease of complexity will be obtained by the application of steerable filters, see section 5.3. Although not obvious, the approximated conjugate Poisson kernel can also be steered by four basis filters.

5.2.3 Isotropy and Coherence

The term $c = \lambda_2 \lambda_1^{-1}$ indicates a measure for the isotropy of the local structure. The extreme cases are $c = 1$ which means that the structure is isotropic and $c = 0$ which means that the structure is i1D.[3] Obviously, this measure of isotropy must be related to the coherence of the structure tensor (3.53). Since the structure tensor yields an estimate of the metric tensor, the λ_k in both tensors can be identified. This is also the reason for using the symbol λ_k^2 instead of λ_k for the eigenvalues of the structure tensor. The following short derivation yields the relationship between coherence and isotropy. In a first step, the two main axes are split into the isotropic part and

[3] Without loss of generality I assume $\lambda_2 < \lambda_1$ in the following.

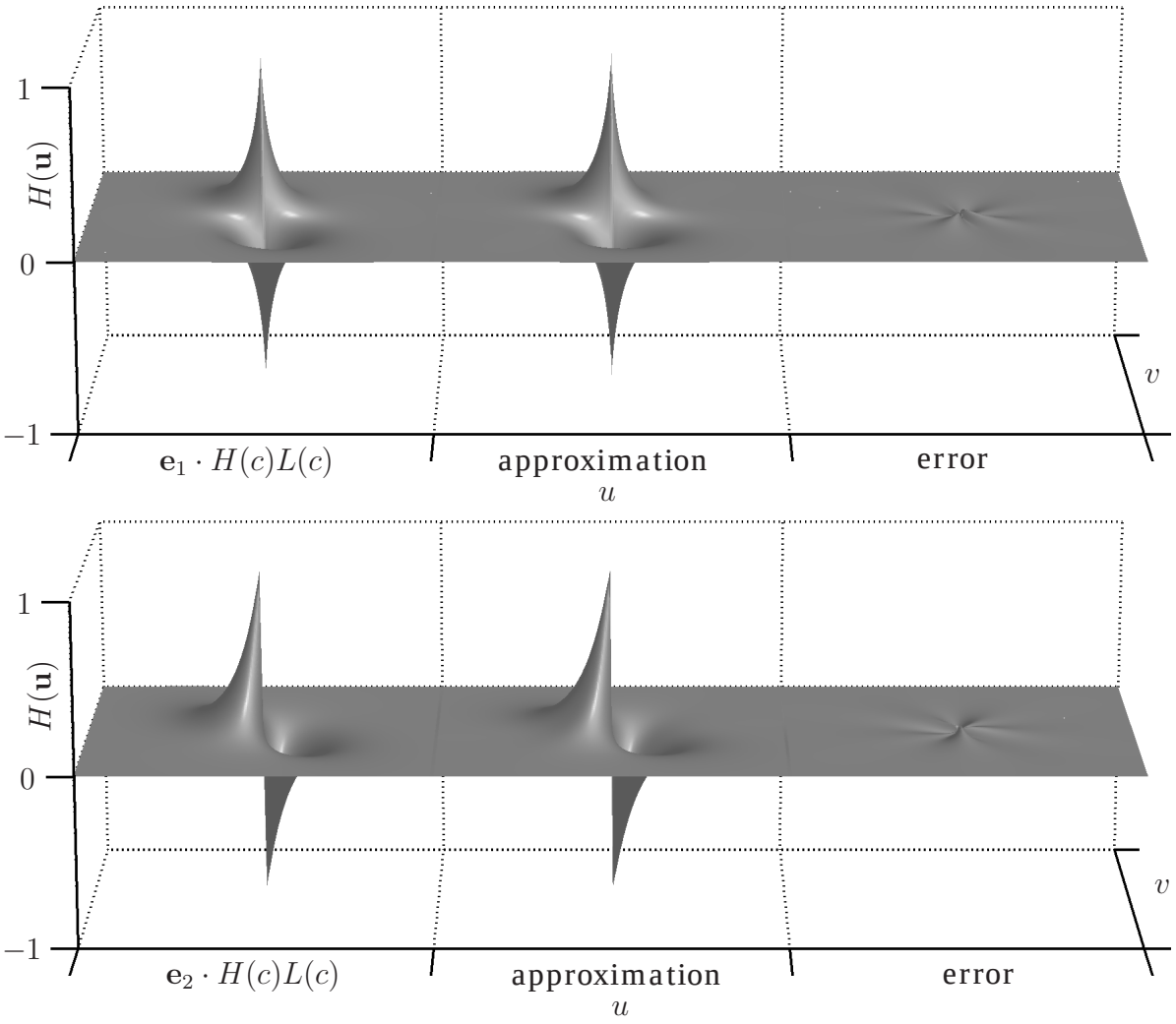

Fig. 5.6: The two components of the anisotropic conjugate Poisson kernel for $c = 2/3$ (left), the approximation by (5.23) (center), and the resulting error functions (right).

the remaining part:

$$\begin{aligned}\lambda_1^2(\mathbf{n}\cdot\mathbf{u})^2 - \lambda_2^2(\mathbf{n}\wedge\mathbf{u})^2 &= \frac{1}{4}(\lambda_1^2(\mathbf{nu}+\mathbf{un})^2 - \lambda_2^2(\mathbf{nu}-\mathbf{un})^2)\\ &= \frac{1}{4}((\lambda_1^2-\lambda_2^2)(\mathbf{nunu}+\mathbf{unun}) + (\lambda_1^2+\lambda_2^2)(\mathbf{nuun}+\mathbf{unnu}))\\ &= \frac{\lambda_1^2+\lambda_2^2}{2}\mathbf{u}^2 + \frac{\lambda_1^2-\lambda_2^2}{2}(\mathbf{nun})\cdot\mathbf{u}\ .\end{aligned}$$

Dividing the whole expression by $(\lambda_1^2+\lambda_2^2)/2$ only affects the total scale s and yields

a new scale $s'' = \sqrt{(\lambda_1^2 + \lambda_2^2)/2}s$. Hence, the final expression

$$\lambda_1^2(\mathbf{n}\cdot\mathbf{u})^2 - \lambda_2^2(\mathbf{n}\wedge\mathbf{u})^2 = \left(\mathbf{u}^2 + \frac{\lambda_1^2-\lambda_2^2}{\lambda_1^2+\lambda_2^2}(\mathbf{nun})\cdot\mathbf{u}\right)\frac{\lambda_1^2+\lambda_2^2}{2} \tag{5.24}$$

is obtained. The first fraction in the expression above is the square root of the coherence according to (3.53). In order to express the coherence by means of the isotropy factor c, use the equalities $\lambda_1^2 = c^{-1}\lambda_1\lambda_2$ and $\lambda_2^2 = c\lambda_1\lambda_2$. The resulting equation reads

$$\frac{\lambda_1^2-\lambda_2^2}{\lambda_1^2+\lambda_2^2} = \frac{c^{-1}-c}{c^{-1}+c} = \frac{1-c^2}{1+c^2}\ , \tag{5.25}$$

which establishes a direct relation between (the square root of) the coherence and the isotropy factor c. It depends on the application which of both measures for the isotropy or the intrinsic dimension is better suited. In figure 5.7 some examples for ellipses and the corresponding isotropy measures can be found.

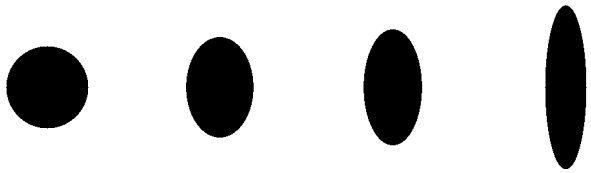

Fig. 5.7: From left to right: ellipses for the isotropies $c = 1$ (circle), $c = 2/3$, $c = 1/2$, and $c = 1/4$. The corresponding coherences according to (5.25) are 0, 0.15, 0.36, and 0.78.

Besides the relation between coherence and isotropy factor, the total scale differs in the two cases by a factor of

$$\frac{s'}{s''} = \sqrt{\frac{2\lambda_1\lambda_2}{\lambda_1^2+\lambda_2^2}} = \sqrt{\frac{2}{c^{-1}+c}}$$

which is just the square root of the ratio between the mixed term and the quadratic terms of the polynomial $(\lambda_1+\lambda_2)^2$. The question is, however, which one of the two terms is better suited to express the scale: the square root of the area of the ellipse created by λ_1 and λ_2 or the sum of the squared axes lengths (see figure 5.1). The answer to this question cannot be given by straightforward calculation or by short reasoning, I will leave it open at this point.

The main result of section 5.2 is that the approximations (5.20) and (5.22) reduce the computational load for calculating the EQF significantly, since the linear combination of one isotropic filter and one oriented anisotropic filter can be used to compute a 1D space of anisotropic filters (depending on c). The original space of

anisotropic filters is 3D (if the relative bandwidth is constant[4]): one dimension for the scale (center frequency) s, one dimension for the isotropy c, and one dimension for the orientation θ_0. The scale can be covered by a pyramid which decreases the computational complexity to $O(N^2)$. By means of the approximations, the space of filters is reduced by one dimension, and hence the complexity for the application of the filters is $O(N^{3/2})$. In the next section, the complexity will be further decreased by applying the method of steerable filters.

5.3 Steering of the Approximation

The idea of the approximations (5.20) and (5.22) is to replace a 1D subspace of filters by linear combinations of basis filters. This idea is quite similar to the idea of steerable filters, introduced in section 3.5. In this section the approach of steerable filters is applied to the orientation depending terms of the EQF approximation. As a result, the computational complexity can be reduced to $O(N)$, and therefore it only differs by a constant factor from the complexity of the SQF computation.

5.3.1 Steering the Poisson Kernel Approximation

Taking a closer look at $L_c(1)$, i.e., the oriented part of the approximation (5.20), the orientation information is contained in the term $\cos(2\alpha)$, where α is the difference between the angular coordinate θ and the orientation of the ellipse θ_0. Obviously, this term can be obtained from a steered spherical harmonic of second order. The aim is now to calculate $L_c(1)\mathbf{F}(\mathbf{u})$ with only one convolution, where $\mathbf{F}(\mathbf{u})$ is the Fourier transform of an arbitrary $\mathbf{e}_3$-valued signal $\mathbf{f}(\mathbf{x})$.

The radial weighting of $L_c(1)$ in (5.20) is given by $|\mathbf{u}| \exp(-2\pi|\mathbf{u}|s')$, and hence the basis function which is needed for steering $L_c(1)\mathbf{F}(\mathbf{u})$ reads in the frequency domain

$$G_2^1(\mathbf{u}) = \frac{\mathbf{u}\mathbf{e}_1\mathbf{u}}{|\mathbf{u}|} \exp(-2\pi|\mathbf{u}|s')\mathbf{F}(\mathbf{u}) \ , \tag{5.26}$$

where the subscript 2 indicates that a second order harmonic is involved and the superscript 1 indicates that the radial weighting is given by the first power of $|\mathbf{u}|$ times the Poisson frequency response.

Since the applied frequency response corresponds to an $\{\mathbf{e}_1, \mathbf{e}_2\}$-valued impulse response (see appendix A.2) and $\mathbf{f}$ is $\mathbf{e}_3$-valued, the spatial representation $g_2(\mathbf{x})$ of $G_2(\mathbf{x})$ is $\{\mathbf{e}_{31}, \mathbf{e}_{23}\}$-valued.

[4] Actually, the bandwidth can also be changed by combining fixed bandwidth filters with different center frequencies. This approach is not considered more in detail, I keep with a fixed decomposition of the radial frequency.

The isotropic part of the approximation, i.e., $L(1)$, is just the standard Poisson kernel with scale s' applied to $\mathbf{f}$, and hence $L(1)\mathbf{F}(\mathbf{u})$ is given in the frequency domain as

$$G_0^0(\mathbf{u}) = \exp(-2\pi|\mathbf{u}|s')\mathbf{F}(\mathbf{u}) \ ,$$

where the subscript 0 indicates that a zeroth order harmonic is involved and the superscript 0 indicates that the Poisson frequency response is weighted by the zeroth power of $|\mathbf{u}|$.

By means of the filter outputs g_0^0 and g_2^1, the anisotropic smoothed signal is approximated by

$$(l(c) * \mathbf{f})(\mathbf{x}) \approx g_0^0(\mathbf{x}) + (c-1)\pi s' \, (\mathbf{n}\mathbf{e}_1\mathbf{n}) \cdot g_2^1(\mathbf{x}) \ , \tag{5.27}$$

where $\mathbf{n} = \cos\theta_0\mathbf{e}_1 + \sin\theta_0\mathbf{e}_2$ denotes the normal orientation vector. In the following I will show that (5.27) is indeed the filter response of (5.20). Although this proof is not complicated to follow, it is not necessary for understanding the remaining part of the thesis.

Considering the inner product $(\mathbf{n}\mathbf{e}_1\mathbf{n}) \cdot g_2^1(\mathbf{x})$ (inner product of a vector and a bivector) in the Fourier domain yields

$$\begin{aligned}
\mathcal{F}_2\{(\mathbf{n}\mathbf{e}_1\mathbf{n}) \cdot g_2^1(\mathbf{x})\}(\mathbf{u}) &= \frac{1}{2}\mathcal{F}_2\{\mathbf{n}\mathbf{e}_1\mathbf{n}g_2^1(\mathbf{x}) - g_2^1(\mathbf{x})\mathbf{n}\mathbf{e}_1\mathbf{n}\}(\mathbf{u}) \\
&= \frac{1}{2}(\mathbf{n}\mathbf{e}_1\mathbf{n}G_2^1(\mathbf{u}) - G_2^1(\mathbf{u})\mathbf{n}\mathbf{e}_1\mathbf{n}) \\
&= \frac{\exp(-2\pi|\mathbf{u}|s')}{2|\mathbf{u}|}(\mathbf{n}\mathbf{e}_1\mathbf{n}\mathbf{u}\mathbf{e}_1\mathbf{u}\mathbf{F}(\mathbf{u}) - \mathbf{u}\mathbf{e}_1\mathbf{u}\mathbf{F}(\mathbf{u})\mathbf{n}\mathbf{e}_1\mathbf{n}) \\
&\quad \text{since } \mathbf{F}(\mathbf{u}) \text{ is } \{\mathbf{e}_3, \mathbf{e}_{12}\}\text{-valued and therefore} \\
&\quad \text{anticommutes with vectors in } \{\mathbf{e}_1, \mathbf{e}_2\} \text{ gives} \\
&= \frac{\exp(-2\pi|\mathbf{u}|s')}{2|\mathbf{u}|}(\mathbf{n}\mathbf{e}_1\mathbf{n}\mathbf{u}\mathbf{e}_1\mathbf{u} + \mathbf{u}\mathbf{e}_1\mathbf{u}\mathbf{n}\mathbf{e}_1\mathbf{n})\mathbf{F}(\mathbf{u}) \\
&= \frac{\exp(-2\pi|\mathbf{u}|s')}{|\mathbf{u}|}(\mathbf{n}\mathbf{e}_1\mathbf{n}) \cdot (\mathbf{u}\mathbf{e}_1\mathbf{u})\,\mathbf{F}(\mathbf{u}) \\
&= |\mathbf{u}|\cos(2\alpha)\exp(-2\pi|\mathbf{u}|s')\mathbf{F}(\mathbf{u}) \ .
\end{aligned}$$

The inner product $(\mathbf{n}\mathbf{e}_1\mathbf{n}) \cdot (\mathbf{u}\mathbf{e}_1\mathbf{u})$ is just the scalar product of the two vectors $\cos(2\theta_0)\mathbf{e}_1 + \sin(2\theta_0)\mathbf{e}_2$ and $\mathbf{u}^2(\cos(2\theta)\mathbf{e}_1 + \sin(2\theta)\mathbf{e}_2)$ and is therefore equivalent to $\mathbf{u}^2\cos(2(\theta - \theta_0)) = \mathbf{u}^2\cos(2\alpha)$. The proof is completed by considering the whole expression (5.27) in the frequency domain which yields the operator output of (5.20):

$$\mathcal{F}_2\{g_0^0(\mathbf{x}) + (c-1)\pi s' \, (\mathbf{n}\mathbf{e}_1\mathbf{n}) \cdot g_2^1(\mathbf{x})\} = (1 + (c-1)\pi s'|\mathbf{u}|\cos(2\alpha))\exp(-2\pi|\mathbf{u}|s')\mathbf{F}(\mathbf{u}) \, .$$

Hence, the application of the anisotropic Poisson kernel can be approximated by means of (5.27) which yields two convolutions, one inner product computation, and

one linear combination. Since the anisotropic bandpass of the EQF is composed by two 'compatible' Poisson kernels, i.e., both kernels have the same isotropy, c, and the same main orientation, the computational complexity is the same for an elliptic bandpass as for the anisotropic Poisson kernel. An implementation based on solely real valued operations (i.e., convolve component-wise) performs three convolutions, three multiplications, and three additions, and is therefore very efficient.

5.3.2 Steering the Conjugate Poisson Kernel Approximation

Steering of the approximation is more complicated in the case of the conjugate Poisson kernel than for the Poisson kernel. The final formulation of the steered approximation requires four different basis filter outputs which are given by

$$G_1^0(\mathbf{u}) = \frac{\mathbf{u}}{|\mathbf{u}|} I_2^{-1} \exp(-2\pi|\mathbf{u}|s')\mathbf{F}(\mathbf{u}) \tag{5.28}$$

$$G_1^1(\mathbf{u}) = \mathbf{u}\, I_2^{-1} \exp(-2\pi|\mathbf{u}|s')\mathbf{F}(\mathbf{u}) \tag{5.29}$$

$$G_3^0(\mathbf{u}) = \frac{\mathbf{u}\mathbf{e}_1\mathbf{u}\mathbf{e}_1\mathbf{u}}{|\mathbf{u}|^3} I_2^{-1} \exp(-2\pi|\mathbf{u}|s')\mathbf{F}(\mathbf{u}) \quad \text{and} \tag{5.30}$$

$$G_3^1(\mathbf{u}) = \frac{\mathbf{u}\mathbf{e}_1\mathbf{u}\mathbf{e}_1\mathbf{u}}{|\mathbf{u}|^2} I_2^{-1} \exp(-2\pi|\mathbf{u}|s')\mathbf{F}(\mathbf{u}) \quad , \tag{5.31}$$

where the subscript indicates the order of the involved spherical harmonic and the superscript indicates the power of $|\mathbf{u}|$. Based on these four basis filter outputs, the output of the anisotropic conjugate Poisson filter is approximated by

$$(h(c) * l(c) * \mathbf{f})(\mathbf{x}) \approx g_1^0(\mathbf{x}) + (c-1)\mathbf{e}_1 \left(\frac{1}{4}(\mathbf{n}\mathbf{e}_1\mathbf{n} g_3^0(\mathbf{x}) - g_1^0(\mathbf{x})\mathbf{n}\mathbf{e}_1\mathbf{n}) + \frac{\pi s'}{2}(g_1^1(\mathbf{x})\mathbf{n}\mathbf{e}_1\mathbf{n} + \mathbf{n}\mathbf{e}_1\mathbf{n} g_3^1(\mathbf{x})) \right) \quad . \tag{5.32}$$

Although not necessary for further reading, the derivation of the equation above helps to understand the behavior of the approximation in the frequency domain.

The proof is based on two equivalences. First, it must be shown that

$$-\frac{\mathbf{e}_1}{4}(g_1^0(\mathbf{x})\mathbf{n}\mathbf{e}_1\mathbf{n} - \mathbf{n}\mathbf{e}_1\mathbf{n} g_3^0(\mathbf{x})) = (h_c(1) * l(1) * \mathbf{f})(\mathbf{x})$$

and second, it must be shown that

$$\frac{\pi s'\mathbf{e}_1}{2}(g_1^1(\mathbf{x})\mathbf{n}\mathbf{e}_1\mathbf{n} + \mathbf{n}\mathbf{e}_1\mathbf{n} g_3^1(\mathbf{x})) = (h(1) * l_c(1) * \mathbf{f})(\mathbf{x}) \quad .$$

By these two identities, the response of (5.23) is obtained by applying the Fourier transform and since $g_1^0(\mathbf{x}) = (h(1) * l(1) * \mathbf{f})(\mathbf{x})$.

Both identities are obtained by straightforward calculations. For the first one, considering the Fourier transform of $h_c(1) * \mathbf{f}$ gives (see (5.21)):

$$\begin{aligned}
\frac{\mathbf{nun} - \mathbf{u}\cos(2\alpha)}{2|\mathbf{u}|} I_2^{-1}\mathbf{F}(\mathbf{u}) &= \frac{|\mathbf{u}|^2\mathbf{ne}_1\mathbf{e}_1\mathbf{un} - \mathbf{u}(\mathbf{ne}_1\mathbf{n})\cdot(\mathbf{ue}_1\mathbf{u})}{2|\mathbf{u}|^3} I_2^{-1}\mathbf{F}(\mathbf{u}) \\
&= \frac{2|\mathbf{u}|^2\mathbf{e}_1\mathbf{une}_1\mathbf{n} - \mathbf{une}_1\mathbf{nue}_1\mathbf{u} - |\mathbf{u}|^2\mathbf{e}_1\mathbf{une}_1\mathbf{n}}{4|\mathbf{u}|^3} I_2^{-1}\mathbf{F}(\mathbf{u}) \\
&= \frac{\mathbf{e}_1\mathbf{une}_1\mathbf{n}}{4|\mathbf{u}|} I_2^{-1}\mathbf{F}(\mathbf{u}) - \frac{\mathbf{e}_1\mathbf{ne}_1\mathbf{nue}_1\mathbf{ue}_1\mathbf{u}}{4|\mathbf{u}|^3} I_2^{-1}\mathbf{F}(\mathbf{u}) \\
&= \frac{\mathbf{e}_1}{4}\left(\frac{\mathbf{u}}{|\mathbf{u}|} I_2^{-1}\mathbf{F}(\mathbf{u})\mathbf{ne}_1\mathbf{n} - \mathbf{ne}_1\mathbf{n}\frac{\mathbf{ue}_1\mathbf{ue}_1\mathbf{u}}{|\mathbf{u}|^3} I_2^{-1}\mathbf{F}(\mathbf{u})\right) \quad .
\end{aligned}$$

Multiplying both sides by $-\exp(-2\pi|\mathbf{u}|s')$ and applying the inverse Fourier transform completes the first proof.

For the second proof, I start with the Fourier transform of the left hand side which gives

$$\begin{aligned}
\frac{\mathbf{e}_1}{2}(G_1^1(\mathbf{u})\mathbf{ne}_1\mathbf{n} + \mathbf{ne}_1\mathbf{n}G_3^1(\mathbf{u})) &= \frac{\mathbf{e}_1}{2}\left(\mathbf{u}\, I_2^{-1}\mathbf{F}(\mathbf{u})\mathbf{ne}_1\mathbf{n} + \mathbf{ne}_1\mathbf{n}\frac{\mathbf{ue}_1\mathbf{ue}_1\mathbf{u}}{\mathbf{u}^2} I_2^{-1}\mathbf{F}(\mathbf{u})\right) L(1) \\
&= \frac{1}{2}\left(\frac{-\mathbf{u}^2\mathbf{e}_1\mathbf{u}I_2^{-1}\mathbf{ne}_1\mathbf{n} + \mathbf{e}_1\mathbf{ne}_1\mathbf{nue}_1\mathbf{ue}_1\mathbf{u}I_2^{-1}}{\mathbf{u}^2}\right) L(1)\mathbf{F}(\mathbf{u}) \\
&= \frac{\mathbf{u}}{|\mathbf{u}|} I_2^{-1}\left(\frac{\mathbf{ue}_1\mathbf{une}_1\mathbf{n} + \mathbf{ne}_1\mathbf{nue}_1\mathbf{u}}{2|\mathbf{u}|}\right) L(1)\mathbf{F}(\mathbf{u}) \\
&= H(1)\left(\frac{(\mathbf{ue}_1\mathbf{u})\cdot(\mathbf{ne}_1\mathbf{n})}{|\mathbf{u}|}\right) L(1)\mathbf{F}(\mathbf{u}) \quad ,
\end{aligned}$$

which is equivalent to $H(1)|\mathbf{u}|\cos(2\alpha)\exp(-2\pi|\mathbf{u}|s')\mathbf{F}(\mathbf{u})$. Multiplying both sides by $\pi s'$ and applying the inverse Fourier transform completes the second proof.

Hence, the combination of the previous two identities yields the Fourier transform of (5.32):

$$\begin{aligned}
&\mathcal{F}_2\left\{g_1^0(\mathbf{x}) + (c-1)\mathbf{e}_1\left(\frac{1}{4}(\mathbf{ne}_1\mathbf{n}g_3^0(\mathbf{x}) - g_1^0(\mathbf{x})\mathbf{ne}_1\mathbf{n}) + \frac{\pi s'}{2}(g_1^1(\mathbf{x})\mathbf{ne}_1\mathbf{n} + \mathbf{ne}_1\mathbf{n}g_3^1(\mathbf{x}))\right)\right\} \\
&\quad = \left(\frac{\mathbf{u}}{|\mathbf{u}|} + (c-1)\left(\frac{\mathbf{u}\cos(2\alpha) - \mathbf{nun}}{2|\mathbf{u}|} + \pi s'\mathbf{u}\cos(2\alpha)\right)\right) I_2^{-1}\exp(-2\pi|\mathbf{u}|s')\mathbf{F}(\mathbf{u}) \quad ,
\end{aligned}$$

which is the operator output of (5.23) in the frequency domain.

The spherical harmonics of order zero to three already occurred in [49] in the context of an i2D generalization of the monogenic signal. In contrast to the issue in the latter paper, the appropriate approximation of the anisotropic (conjugate) Poisson kernel also requires the $|\mathbf{u}|$-weighted spherical harmonics of order one to three. The impulse responses of all used spherical harmonics are derived in appendix A.2.

Whereas the isotropic Poisson kernel yields an isotropic linear scale-space (see section 4.5), the anisotropic Poisson kernel obviously allows to smooth the signal in an adaptive, anisotropic way. As I have shown, the oriented anisotropy can efficiently be adopted according to the local signal structure, due to the steered first order approximation of the anisotropic Poisson kernel. The resulting filter scheme has to be compared to methods of anisotropic diffusion (see e.g. [116, 134, 142]). However, it is still unknown how to estimate the local isotropy in order to have an adaptive behavior. The estimation of this local isotropy is part of the i2D extension of the monogenic signal which is developed in the remaining part of this chapter.

5.4 The Isotropic Analytic Signal for i2D Signals

In this section a 2D signal model is introduced which is motivated by considerations about sampling and the structure tensor. The 2D model is based on two i1D signals which are perpendicular, i.e., their orientation vectors are orthogonal. The behavior of the monogenic signal for this kind of signals is surely suboptimal although better than the results which are obtained from classical approaches (see section 1.2).

Hence, the main topic of this section is to extend the monogenic signal such that it allows to analyze 2D signals. Based on the introduced signal model and the idea of the quaternionic analytic signal [23], a local rotation invariant signal decomposition is developed. In order to obtain an isotropic approach, the spherical harmonics of order zero to three are applied as basis filters, yielding a 7D filter response which is embedded in $\mathbb{R}_3$. Steering these filter responses and subsequent linear combinations yield a rotation invariant decomposition of the signal. From these components a rich set of features can be extracted (see section 5.5). The 7D multivector together with its generalized 'phase model' is named the *structure multivector* (SMV).

5.4.1 The 2D Signal Model for i2D Signals

As mentioned in section 1.1, the space of 2D symmetries is infinite times larger than the space of 1D symmetries and it is impossible to cover all 2D symmetries with a finite number of basis filters. Therefore, restricting the 2D symmetries to a subset of all symmetries is necessary, and furthermore it is also reasonable since the angular behavior of a signal can be considered to be *band limited*. This fact arises from a simple estimation of the maximal frequency of the angular coordinate. For a signal of size $N \times N$ the angular coordinate is finest sampled on the boundary, i.e., the minimal angular distance between two samples is $2/(N-1)$ (in radians, obtained by the inverse of the radius). Consequently, it is not reasonable to have more than

$2\pi/(2/(N-1)) = (N-1)\pi$ samples[5] on the full circle, and hence the highest occurring frequency[6] is less or equal to $(N-1)\pi/2$. By such a consideration, the number of 2D symmetries is implicitly reduced to a finite number, since the (here considered) 2D symmetries are composed of 1D symmetries with different angles.

To be more concrete, the band limitation of the frequency of the angular coordinate is depending on the radius and therefore spatially variant. Hence, creating filters with a certain angular frequency component[7] requires that the amplitude of the filter is zero in a region about its origin. Otherwise *aliasing* occurs, as it is the case for the Siemens-star (the greater its angular frequency, the larger the area of aliasing). If a filter is supposed to be local (i.e., the coefficient adjacent to the origin are nonzero) and performs an angular symmetry decomposition, this automatically bounds the angular frequency to be less or equal to three if the bound shall be independent of the orientation to the grid (see figure 5.8). Hence, it is not reasonable to combine a local filter with angular decompositions of orders higher than three.

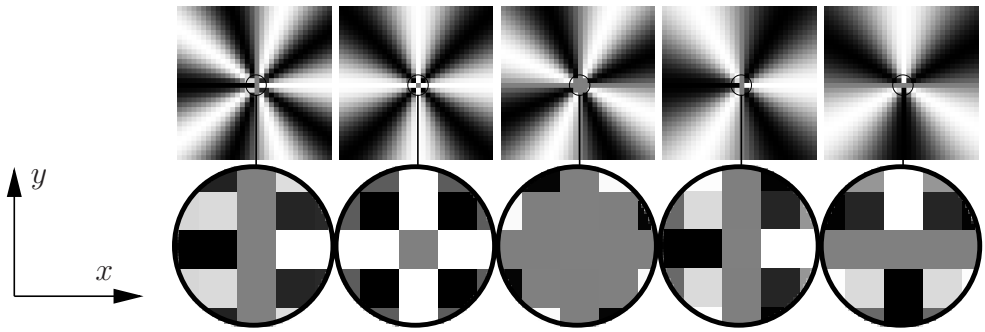

Fig. 5.8: Angular aliasing depending on the radius. From left to right: For a 3×3 area, the angular frequency of value five is too high. An angular frequency of value four does not show aliasing for a 'compatible' orientation, but it does for an 'incompatible' orientation. An angular frequency of value three does not show aliasing in any orientation (fourth and fifth image).

If the angular decomposition is considered by means of a Fourier series, the spherical harmonics of order zero to three correspond to four equidistant sampling points of the angular coordinate. Therefore, the following signal model is introduced. A 2D signal according to this model consists of two i1D signals with perpendicular orientations, i.e.,

$$\mathbf{f}(\mathbf{x}) = \mathbf{f}_1(\mathbf{n}(\mathbf{x}) \cdot \mathbf{x}) + \mathbf{f}_2(\mathbf{n}(\mathbf{x})^* \cdot \mathbf{x}) \ , \tag{5.33}$$

[5] Note that the number of samples on the circle differ from the number of samples on the rectangle boundary. Assuming e.g. $N = 2$ yields π samples. The number of samples on the boundary is 4.

[6] Considering the example in footnote 5 again, the upper bound for the frequency is two and the highest occurring frequency has to be less than this, hence one.

[7] Note that the angular frequency is invariant under the Fourier transform, see (A.1).

where $\mathbf{f}_1$ and $\mathbf{f}_2$ are 1D signals, $\mathbf{n}(\mathbf{x})$ is an arbitrary unit vector, and $\mathbf{n}(\mathbf{x})^* = \mathbf{n}(\mathbf{x})I_2^{-1}$ is its dual in the image plane[8]. An example for the proposed signal model are the basis function of the QFT [21], but in contrast to the latter, the new model is not attached to the coordinate system.

Any arbitrary 2D signal can be approximated by a 2D signal according to the proposed model. If the correct main orientation of the original signal is estimated by $\mathbf{n}$, the approximation is optimal with respect to the L^2-norm due to the optimality of Fourier series. The reduction of the local structure to two perpendicular components is also related to the derivation of the structure tensor which also reflects the main orientation and the component perpendicular to the latter. Further extensions of the i2D signal modeling would yield a better approximation in theory, but in practice the angular aliasing does not allow further terms.

The restriction to two perpendicular i1D signals can be interpreted in the Radon domain as two single lines. However, this is only correct in a local context. Basically, the Radon transform is a global transform, in contrast to the signal model which is local. This locality implies an adaption of the orientation to the signal. Hence, the approximation is much better than it is supposed to be at a first glance. It is important to understand the consequences of the locality of the new model. Consider for example a corner with an angle much smaller than $\pi/2$ (see figure 5.9), hence not being in the class of the modeled i2D signals. By locally choosing the orientation of the coordinate system and the ratio of the two i1D signals, the original signal is approximated very well, as it can be seen in the schematic drawing in figure 5.9 on the right.

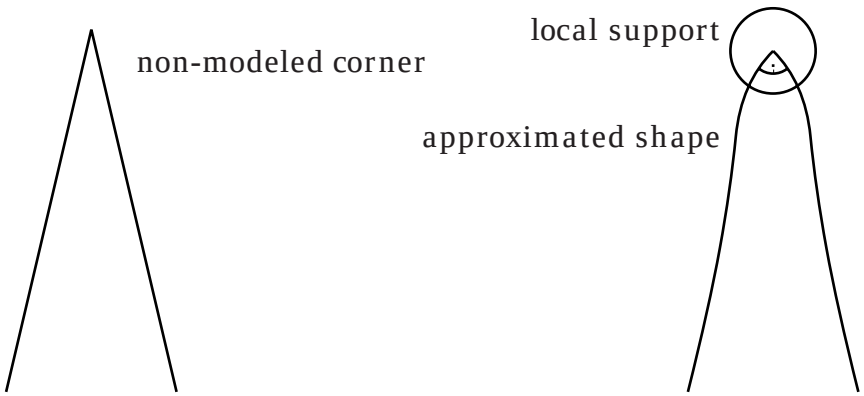

Fig. 5.9: Corner which is not contained in the 2D signal model (left) and its approximation by locally modeled functions (right). Inside the local support (circle) the class of functions is restricted to consist of perpendicular components. In the environment of the circle different orientations can be used.

[8] In this chapter the dual is always formed with respect to the image plane, i.e., $M^* = MI_2^{-1}$.

5.4.2 The Error of the Monogenic Signal for i2D Signals

The monogenic signal and its phase approach are designed for i1D signals, i.e., if i2D signals are considered, certain distortions occur. This is the case for the egg-tray signal (see section 1.1), since it is an i2D signal. The structures in this special signal are perpendicular, and hence according to (4.34) a Riesz component is obtained which is not the correct quadrature component. The obvious results are amplitude and phase distortions, see figure 5.10. Although the phase distortions are much

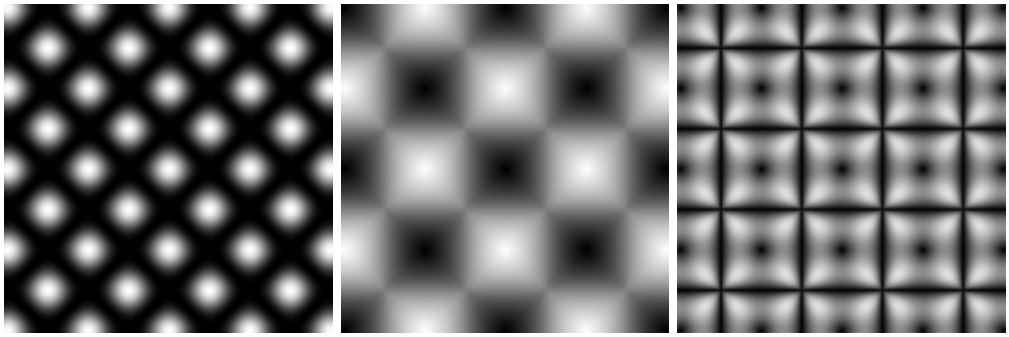

Fig. 5.10: From left to right: Amplitude distortions, phase image, and phase distortions of the monogenic signal applied to the egg-tray signal. The amplitude distortions are visualized by the relative square error in a range from zero to one. The phase image is similar, but better than that of the adaptive partial Hilbert transform (see figure 1.2). The phase distortions are scaled by ten (i.e., black means zero and white means $\pi/10$).

weaker than those of the classical approaches, the question is whether it is possible to combine the ideas of the quaternionic analytic signal and of the monogenic signal in order to have an optimal, isotropic result.

The key to a correct phase estimation is a constant amplitude estimation, i.e., the local amplitude should be independent of the orientation and of the intrinsic dimension. For i1D signals the amplitude of the monogenic signal is constant, but in the case of the egg-tray signal, i.e., for i2D signals, it varies. The error which occurs can directly be calculated. The local energy of the egg-tray signal is constantly two, and therefore the energy error of its monogenic signal is given by

$$\begin{aligned}
|\mathbf{f}_{\mathrm{M}}|^2 - 2 &= (\cos(x+y) + \cos(x-y))^2 + (\sqrt{1/2}\sin(x+y) + \sqrt{1/2}\sin(x-y))^2 \\
&\quad + (\sqrt{1/2}\sin(x+y) - \sqrt{1/2}\sin(x-y))^2 - 2 \\
&= \cos^2(x+y) + 2\cos(x+y)\cos(x-y) + \cos^2(x-y) \\
&\quad + \sin^2(x+y) + \sin^2(x-y) - 2 \\
&= 2\cos(x+y)\cos(x-y) \quad .
\end{aligned}$$

In order to compensate this error, another component must be introduced, such that the term $-2\cos(x+y)\cos(x-y)$ is obtained. This mixed term occurs in $(\cos(x+y)-\cos(x-y))^2$, and therefore a frequency response must be introduced which is negative in the second and fourth quadrant, since $\mathcal{F}_2\{\cos(x-y)\} = \frac{1}{2}(\delta_0(\mathbf{u}-\mathbf{e}_1+\mathbf{e}_2)+\delta_0(\mathbf{u}+\mathbf{e}_1-\mathbf{e}_2))$. Such a frequency response appears as the odd-odd symmetry of the QFT [22], but in contrast to that QFT component, the new frequency response should not be attached to the coordinate system.

Although it seems that adding an odd-odd component is sufficient to obtain a constant amplitude and a correct phase, this remains only true unless the assumed signal model is violated. If the signal contains components which do not lie on either the line of main orientation or on its perpendicular line, amplitude and phase distortions are introduced. In order to obtain a robust estimate of the amplitude and the phase, the generalized analytic signal must contain angular envelopes which fulfill two conditions. First, the envelopes for the even part and the odd part must be the same, and second, the envelope must realize some kind of angular distance measure, i.e., it must monotonically decrease with the angular distance. These two conditions are discussed in the subsequent two sections.

5.4.3 Spectral Decomposition According to Symmetries

In order to obtain the complete phase decomposition at each position, the i2D signal must be separated according to the modeling into its components. Instead of treating arbitrary 2D signals, it is sufficient to consider the basis functions which correspond to the introduced signal model. Similar as in [17], they consist of products of trigonometric functions, but in contrast to the QFT they are not fixed to the coordinate axes. An approach which is appropriate to these basis functions should allow to switch between them, or equivalently, to switch to exponential functions. In 1D the Hilbert transform switches between sine and cosine functions and the analytic signal corresponds to the complex exponential.

First, the products of two perpendicular trigonometric functions yield four basis functions, where $\mathbf{m} = \frac{\mathbf{n}+\mathbf{n}^*}{2}$ and $\mathbf{m}^\perp = \mathbf{m}I_2 = \frac{\mathbf{n}-\mathbf{n}^*}{2}$:

$$\psi_1(\mathbf{x};\mathbf{n}) = \cos(\mathbf{m}\cdot\mathbf{x})\cos(\mathbf{m}^\perp\cdot\mathbf{x}) = \cos(\mathbf{m}\cdot\mathbf{x})\cos((\mathbf{m}\wedge\mathbf{x})^*) \tag{5.34}$$

$$\psi_2(\mathbf{x};\mathbf{n}) = I_3\cos(\mathbf{m}\cdot\mathbf{x})\sin(\mathbf{m}^\perp\cdot\mathbf{x}) = I_3\cos(\mathbf{m}\cdot\mathbf{x})\sin((\mathbf{m}\wedge\mathbf{x})^*) \tag{5.35}$$

$$\psi_3(\mathbf{x};\mathbf{n}) = I_3\sin(\mathbf{m}\cdot\mathbf{x})\cos(\mathbf{m}^\perp\cdot\mathbf{x}) = I_3\sin(\mathbf{m}\cdot\mathbf{x})\cos((\mathbf{m}\wedge\mathbf{x})^*) \tag{5.36}$$

$$\psi_4(\mathbf{x};\mathbf{n}) = -\sin(\mathbf{m}\cdot\mathbf{x})\sin(\mathbf{m}^\perp\cdot\mathbf{x}) = -\sin(\mathbf{m}\cdot\mathbf{x})\sin((\mathbf{m}\wedge\mathbf{x})^*) \tag{5.37}$$

where $\mathbf{n}$ is an arbitrary vector and $\mathbf{n}^* = \mathbf{n}I_2^{-1}$ is its dual vector. The vectors $\mathbf{m}$ and $\mathbf{m}^\perp$ establish a rotated, orthogonal coordinate system, which replaces the standard uv-coordinate system of the QFT. According to the addition theorems of trigonometric functions (see e.g. [16]), these products can be transformed into sums of trigono-

metric functions, which shows that $\psi_1, \ldots, \psi_4$ are actually functions according to the 2D signal model:

$$\psi_1(\mathbf{x};\mathbf{n}) = \frac{1}{2}(\cos(\mathbf{n}\cdot\mathbf{x}) + \cos(\mathbf{n}^*\cdot\mathbf{x})) \tag{5.38}$$

$$\psi_2(\mathbf{x};\mathbf{n}) = \frac{I_3}{2}(\sin(\mathbf{n}\cdot\mathbf{x}) - \sin(\mathbf{n}^*\cdot\mathbf{x})) \tag{5.39}$$

$$\psi_3(\mathbf{x};\mathbf{n}) = \frac{I_3}{2}(\sin(\mathbf{n}\cdot\mathbf{x}) + \sin(\mathbf{n}^*\cdot\mathbf{x})) \tag{5.40}$$

$$\psi_4(\mathbf{x};\mathbf{n}) = \frac{1}{2}(\cos(\mathbf{n}\cdot\mathbf{x}) - \cos(\mathbf{n}^*\cdot\mathbf{x})) \ . \tag{5.41}$$

Hence, the Fourier transforms of $\psi_1, \ldots, \psi_4$ read

$$\Psi_1(\mathbf{u};\mathbf{n}) = \frac{1}{4}(\delta_0(\mathbf{u}-\mathbf{n}) + \delta_0(\mathbf{u}+\mathbf{n}) + \delta_0(\mathbf{u}-\mathbf{n}^*) + \delta_0(\mathbf{u}+\mathbf{n}^*)) \tag{5.42}$$

$$\Psi_2(\mathbf{u};\mathbf{n}) = \frac{1}{4}(\delta_0(\mathbf{u}-\mathbf{n}) - \delta_0(\mathbf{u}+\mathbf{n}) - \delta_0(\mathbf{u}-\mathbf{n}^*) + \delta_0(\mathbf{u}+\mathbf{n}^*)) \tag{5.43}$$

$$\Psi_3(\mathbf{u};\mathbf{n}) = \frac{1}{4}(\delta_0(\mathbf{u}-\mathbf{n}) - \delta_0(\mathbf{u}+\mathbf{n}) + \delta_0(\mathbf{u}-\mathbf{n}^*) - \delta_0(\mathbf{u}+\mathbf{n}^*)) \tag{5.44}$$

$$\Psi_4(\mathbf{u};\mathbf{n}) = \frac{1}{4}(\delta_0(\mathbf{u}-\mathbf{n}) + \delta_0(\mathbf{u}+\mathbf{n}) - \delta_0(\mathbf{u}-\mathbf{n}^*) - \delta_0(\mathbf{u}+\mathbf{n}^*)) \ , \tag{5.45}$$

see also figure 5.11.

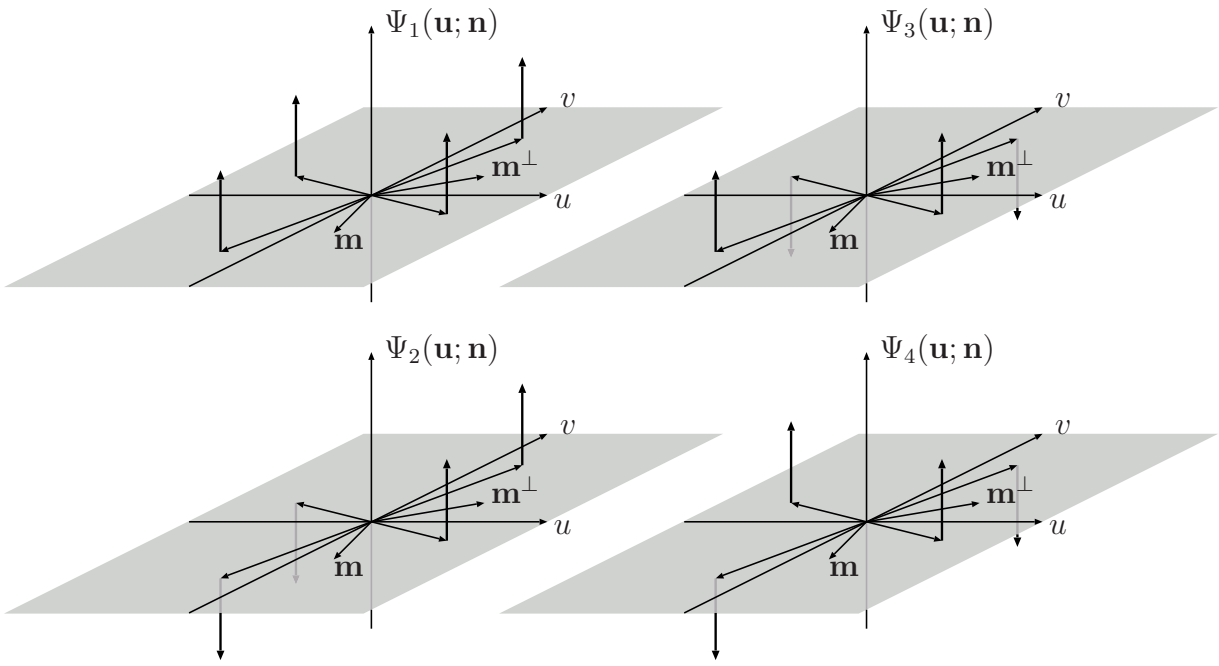

Fig. 5.11: The basis functions $\Psi_1, \ldots, \Psi_4$ in the Fourier domain.

Taking a closer look at the four basis functions, they are just the coordinate system independent equivalents of the QFT basis functions. By substituting $\mathbf{m} = \mathbf{e}_1$ (i.e., $\mathbf{n} = \mathbf{e}_1 + \mathbf{e}_2$), the basis functions are basically identical and their assignments to each other are given in table 5.1.

symmetry	even-even	even-odd	odd-even	odd-odd
basis function	$\Psi_1(\mathbf{u}; \mathbf{e}_1 + \mathbf{e}_2)$	$\Psi_2(\mathbf{u}; \mathbf{e}_1 + \mathbf{e}_2)$	$\Psi_3(\mathbf{u}; \mathbf{e}_1 + \mathbf{e}_2)$	$\Psi_4(\mathbf{u}; \mathbf{e}_1 + \mathbf{e}_2)$

Tab. 5.1: Assignments between the QFT symmetry relations and the new coordinate free basis functions.

From the basis functions $\psi_1, \ldots, \psi_4$ any of the involved trigonometric functions can be extracted, e.g., $\cos(\mathbf{n}^* \cdot \mathbf{x}) = \psi_1(\mathbf{x}; \mathbf{n}) - \psi_4(\mathbf{x}; \mathbf{n})$. Furthermore, this basis can also be used to extract 'complex' harmonics, e.g., $\exp(\mathbf{n} \cdot \mathbf{x}\, I_3) = \psi_1(\mathbf{x}; \mathbf{n}) + \psi_2(\mathbf{x}; \mathbf{n}) + \psi_3(\mathbf{x}; \mathbf{n}) + \psi_4(\mathbf{x}; \mathbf{n})$. Hence, by means of these basis functions, any 2D signal according to the proposed model can be decomposed into its two i1D signals and the corresponding Hilbert transforms. This decomposition is discussed in section 5.5 in detail.

5.4.4 Rotation Invariant 2D Decomposition

The signal decomposition according to the basis functions in figure 5.11 corresponds to knowing quadrature pairs in 1D, i.e., knowing that cosine and sine are in quadrature. The key to a generalized 2D analytic signal is the set of relations between the four basis functions $\Psi_1, \ldots, \Psi_4$, which corresponds to the Hilbert transform in 1D. Moreover, a rotation invariant, odd relationship which is obtained from, e.g., Ψ_1 and Ψ_2, cannot be scalar valued. The reason is quite obvious after the discussion about the Riesz transform in chapter 4.

The derivation of the relationships between the basis functions must be consistent with the assumed signal model, i.e., the resulting operator output should contain the two monogenic signals which correspond to the involved perpendicular i1D signals. Therefore, the 2D signal must be split into its two i1D parts and the Riesz transform of each of them must be evaluated. The decomposition has to be robust (see section 5.4.1), which is achieved by choosing an appropriate angular windowing function.

For the following considerations assume that α is the angle between the main orientation $\theta_0 = \langle \log(\mathbf{e}_1 \mathbf{n}) \rangle_2$ and the frequency vector, i.e., $\alpha = \langle \log(\mathbf{n}\mathbf{u}) \rangle_2$. The estimation of the main orientation will be discussed in section 5.5, just assume for now that it is known. According to the discussion in section 5.4.2, a frequency response

with odd-odd symmetry must be introduced in order to compensate the amplitude error. Considering the spherical harmonics of order zero to three, the only possible choice which fulfills the odd-odd symmetry is $\cos(2\alpha)$. Using the basic identities

$$W_1(\mathbf{u}) = \cos^2(\alpha) = \frac{1+\cos(2\alpha)}{2} \text{ and } W_2(\mathbf{u}) = \sin^2(\alpha) = \frac{1-\cos(2\alpha)}{2} \ , \tag{5.46}$$

two *angular windowing functions* which yield the perpendicular i1D components of the 2D signal are obtained. Those angular windowing functions are also optimal with respect to robustness, since they provide a measure of the angular distance which is the same as the one of the structure tensor.

The necessary frequency response $\cos(2\alpha)$ can be obtained by steering[9] and projecting the spinor field, which is basically given as the second order spherical harmonic[10]:

$$H_2^2(\mathbf{u}) = \frac{\mathbf{e}_1\mathbf{u}\mathbf{e}_1\mathbf{u}}{\mathbf{u}^2} \ . \tag{5.47}$$

This spinor field rotates $\mathbf{e}_1$ into the vectors which are obtained by reflecting $\mathbf{e}_1$ at the frequency vectors, i.e., $\mathbf{e}_1 H_2^2(\mathbf{u}) = \mathbf{u}\mathbf{e}_1\mathbf{u}$ (see figure 5.12). This reflection at the vector field directly corresponds to the double angle representation of the frequency vectors [61], and therefore it is called the *double angle vector field*. In this field vectors at perpendicular frequency vectors (e.g. at $\mathbf{u}_1 = \mathbf{e}_1$ and $\mathbf{u}_2 = \mathbf{e}_2$) are antiparallel, and hence the odd-odd symmetry is achieved. For any main orientation θ_0 the vector field can be rotated such that the projection onto $\mathbf{e}_1$ yields the desired frequency response:

$$(\mathbf{e}_1\mathbf{n}\mathbf{e}_1\mathbf{n}\mathbf{u}\mathbf{e}_1\mathbf{u})\cdot\mathbf{e}_1 = (\mathbf{n}\mathbf{e}_1\mathbf{n})\cdot(\mathbf{u}\mathbf{e}_1\mathbf{u}) = \cos(2\alpha) \ , \tag{5.48}$$

which is the inner product of the double angle orientation vector and the double angle frequency vector field. Plugging (5.48) into (5.46) yields the final expressions for W_1 and W_2:

$$W_1(\mathbf{u}) = \frac{1+(\mathbf{n}\mathbf{e}_1\mathbf{n})\cdot(\mathbf{e}_1 H_2^2(\mathbf{u}))}{2} = \frac{2+(\mathbf{n}\mathbf{e}_1)^2 H_2^2(\mathbf{u}) + (\mathbf{e}_1\mathbf{n})^2\tilde{H}_2^2(\mathbf{u})}{4} \tag{5.49}$$

$$W_2(\mathbf{u}) = \frac{1-(\mathbf{n}\mathbf{e}_1\mathbf{n})\cdot(\mathbf{e}_1 H_2^2(\mathbf{u}))}{2} = \frac{2-(\mathbf{n}\mathbf{e}_1)^2 H_2^2(\mathbf{u}) - (\mathbf{e}_1\mathbf{n})^2\tilde{H}_2^2(\mathbf{u})}{4} \ . \tag{5.50}$$

After splitting the signal into its two i1D parts, the corresponding Riesz transforms have to be evaluated in order to obtain the two i1D phases. Since the main orientation is known beforehand, it is also necessary to steer the Riesz transform.

[9] For the sake of simplicity, I assume that the main orientation is constant, which allows to apply the steering operations in the Fourier domain. The general case, i.e., non-constant main orientation, is obtained by steering in the spatial domain solely, see next section.

[10] In order to obtain a consistent and compact notation, the spherical harmonics of order k are denoted by H_2^k. Hence, the Riesz transform is denoted by H_2^1 in the subsequent sections.

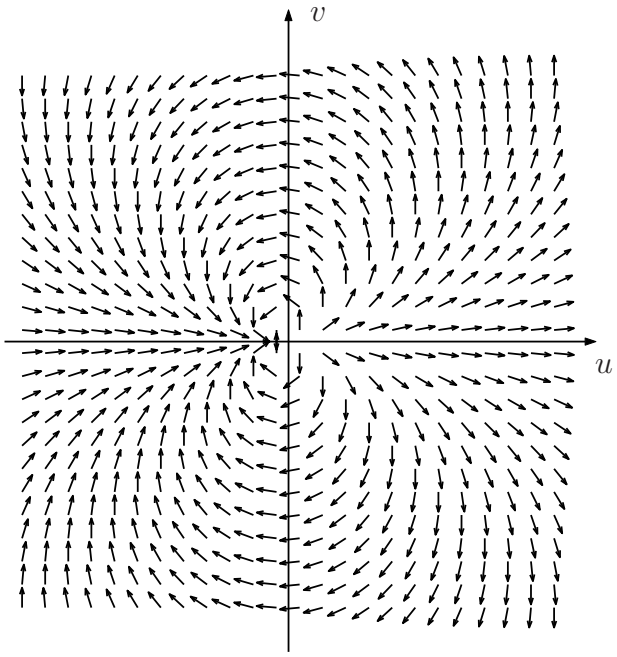

Fig. 5.12: New frequency response which corrects the energy error of the monogenic signal. To visualize the spinor field as a vector field, it is applied from the right to the vector $\mathbf{e}_1$.

The resulting odd functions are formed by a linear combinations of the Riesz frequency response and the *third order spherical harmonic*. The latter is defined by the frequency response

$$H_2^3(\mathbf{u}) = \frac{\mathbf{u}\mathbf{e}_1\mathbf{u}\mathbf{e}_1\mathbf{u}}{|\mathbf{u}|^3} I_2^{-1} \quad , \tag{5.51}$$

see figure 5.13.

Using the third order harmonic, the steered Riesz transform of $W_1(\mathbf{u})$ is given by

$$\begin{aligned} W_3'(\mathbf{u}) &= \mathbf{e}_1\mathbf{n}H_2^1(\mathbf{u})W_1(\mathbf{u}) \\ &= \frac{2\mathbf{e}_1\mathbf{n}H_2^1(\mathbf{u}) + \mathbf{e}_1\mathbf{n}H_2^1(\mathbf{u})H_2^2(\mathbf{u})(\mathbf{n}\mathbf{e}_1)^2 + \mathbf{e}_1\mathbf{n}H_2^1(\mathbf{u})\tilde{H}_2^2(\mathbf{u})(\mathbf{e}_1\mathbf{n})^2}{4} \\ &= \frac{2\mathbf{e}_1\mathbf{n}H_2^1(\mathbf{u}) + (\mathbf{e}_1\mathbf{n})^3 H_2^3(\mathbf{u}) - \mathbf{n}H_2^1(\mathbf{u})\mathbf{e}_1}{4} \quad . \end{aligned} \tag{5.52}$$

By a similar calculation, the steered Riesz transform of $W_2(\mathbf{u})$ is obtained to be (note that the orientation is now given by $\theta_0 + \pi/2$, i.e., the steering term reads $I_2\mathbf{e}_1\mathbf{n}^*$):

$$W_4'(\mathbf{u}) = I_2 \frac{2\mathbf{e}_1\mathbf{n}H_2^1(\mathbf{u}) - (\mathbf{e}_1\mathbf{n})^3 H_2^3(\mathbf{u}) + \mathbf{n}H_2^1(\mathbf{u})\mathbf{e}_1}{4} \quad . \tag{5.53}$$

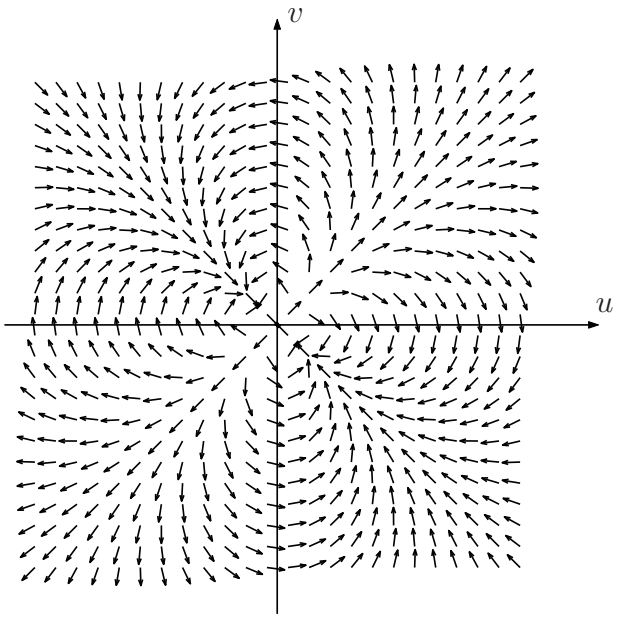

Fig. 5.13: Third order spherical harmonic visualized as a vector field.

Actually, the steering operations which are applied above are already part of the i2D 'phase' model. However, it is hard to motivate the filter derivation without looking ahead to section 5.5, where the responses of the derived filters are decomposed in order to obtain several features of the signal. The main topic of this section has been to motivate the use of the spherical harmonics of order zero to three. The filter responses of these four operators (i.e., before steering) yield seven components altogether.[11] These components are combined as one multivector which can be considered as the *i2D generalization of the analytic signal*. For an $\mathbf{e}_3$-valued signal $\mathbf{f}(\mathbf{x})$ this analytic signal is given by the $\mathbb{R}_3$-multivector

$$M_{\mathrm{S}}(\mathbf{x}) = \mathbf{f}(\mathbf{x}) + (h_2^1 * \mathbf{f})(\mathbf{x}) + \mathbf{e}_3(h_2^2 * \mathbf{f})(\mathbf{x}) + \mathbf{e}_3(h_2^3 * \mathbf{f})(\mathbf{x}) \ . \tag{5.54}$$

Note that the monogenic signal is obtained by neglecting the even-order terms, $\mathbf{f}_{\mathrm{M}}(\mathbf{x}) = \langle M_{\mathrm{S}} \rangle_1$ ('backwards compatibility'). Furthermore, the components are not all weighted equally in the postprocessing as it is the case for the analytic signal and the monogenic signal.

The previous considerations have shown that this i2D analytic signal performs a generalized quadrature decomposition. The postprocessing, which is the topic of the subsequent section, turns the filter responses into a rich set of features which fulfill the invariance – equivariance requirement.

[11] Each of the spherical harmonic responses has two components except for the zeroth.

5.5 The Structure Multivector

In section 5.4 a model for i2D signals has been proposed and a signal decomposition into two i1D signals has been developed. So far, this corresponds to the definition of the analytic signal in terms of the signal and its Hilbert transform or to the monogenic signal in terms of the signal and its Riesz transform. However, the main idea of these approaches is to change the representation such that an amplitude-phase decomposition is obtained. Obviously, a simple change to polar coordinates is not sufficient for the new generalization. A decomposition according to the signal model is developed and it is illustrated by some examples. Similar to the idea of quadrature filters, the new approach, which is called the *structure multivector*, is combined with radial bandpass filters in practice.

5.5.1 Steering by the Local Orientation

As already mentioned in section 5.4.4, steering the basis filter responses becomes crucial in the context of the proposed i2D approach. Fortunately, the local orientation, which is needed for the steering, can directly be estimated from the basis filter responses. The necessary orientation information is even reduced to a 'semi-orientation' which means that it is sufficient to know the orientation in the interval $[0, \pi/2)$. The reason for this reduction is that the 2D signal model includes two perpendicular unit vectors, one of them representing the *main orientation* (or *major orientation*), the other one representing the *minor orientation*. For the steering operation it is sufficient to know either of both, and hence it is sufficient to know the one in the first half of the orientation interval.

Each spherical harmonic of order $k \neq 0$ allows to estimate the orientation of a signal in the interval $[0, \pi/k)$. This is directly obtained by multiplying the frequency response of a spherical harmonic by an impulse line of a certain orientation, i.e., by the spectrum of an i1D signal. The resulting signal in the spatial domain contains some factor of the form $(\mathbf{e}_1\mathbf{n})^k$ which directly allows to estimate the local orientation. Since the sign of the spectrum is basically unknown, one can only estimate the argument of $\pm(\mathbf{e}_1\mathbf{n})^k$, i.e., $\arg((\mathbf{e}_1\mathbf{n})^k) \in [0, \pi)$, and therefore $\arg(\mathbf{e}_1\mathbf{n}) \in [0, \pi/k)$.[12] In section 6.1 a detailed discussion about optimal orientation estimation under the presence of noise is presented with the main result that a certain combination of spherical harmonics gives a good estimate with low computational load. This method is applied in the sequel. For convenience, the following notation is used

$$M_{\mathrm{S}} = M_0 + M_1\mathbf{e}_1 + M_2\mathbf{e}_2 + M_3\mathbf{e}_3 + M_{23}\mathbf{e}_{23} + M_{31}\mathbf{e}_{31} + M_{12}\mathbf{e}_{12} \ , \tag{5.55}$$

[12] Actually, the argument of a spinor is a bivector, hence $\arg(\cdot) \in [\cdot, \cdot)$ means $\arg(\cdot)I_2^{-1} \in [\cdot, \cdot)$.

and the spatial argument of $M_{\mathrm{S}}(\mathbf{x})$ is omitted. The formula for estimating the semi-orientation $\theta_{\mathrm{e}} \in [0, \pi/2)$ reads as follows:

$$\theta_{\mathrm{e}} = \frac{1}{4} \arg((M_0 + M_{12}I_2)^2 + (M_{31} - M_{23}I_2)(M_1 + M_2 I_2)) \ . \tag{5.56}$$

This equation can be verified by calculating each component of the multivector M_{S} for an ideal i1D signal with a fixed, but arbitrary orientation, e.g. $\cos(\mathbf{n} \cdot \mathbf{x})$ (see section 6.1.3). The orientation θ_{e} yields the orientation vector $\mathbf{n} = \cos\theta_{\mathrm{e}}\mathbf{e}_1 + \sin\theta_{\mathrm{e}}\mathbf{e}_2$ which will be used rather than the orientation itself.

The estimated angle is obtained from a *four-fold angle representation*, since both addends in (5.56) are spherical harmonics of order four. In contrast to the remark about the orientation estimation by one harmonic solely, the sign of the signal cancels out in this case and a negative sign of the addends in (5.56) means a different angle. Hence, the angle is not in $[0, \pi/4)$, but in $[0, \pi/2)$, such that the estimated orientation is given in the first quadrant. A second orientation is obtained by implicitly adding $\pi/2$ (see also figure 5.11). These two orientations are the local orientations of the two i1D signals forming the modeled 2D signal. The local orientation which belongs to the stronger i1D signal is called the major orientation and defines the main orientation.

The orientation in the first quadrant can be used to steer the basis filter responses yielding the operator outputs of $\mathcal{W}_1$, $\mathcal{W}_2$, $\mathcal{W}_3'$, and $\mathcal{W}_4'$, see section 5.4.4. These four outputs can be combined to form two monogenic signals. However, the phase model of the monogenic signal includes information about the local orientation, which is not reasonable in this case since the orientation is already known. Hence, the filter responses of $\mathcal{W}_3'$ and $\mathcal{W}_4'$ must be projected onto $-\mathbf{e}_2$.[13] This yields the two frequency responses

$$\begin{aligned} W_3(\mathbf{u}) = -W_3'(\mathbf{u}) \cdot \mathbf{e}_2 &= \frac{3\langle \mathbf{n} H_2^1(\mathbf{u}) I_2 \rangle_0 + \langle \mathbf{n}\mathbf{e}_1\mathbf{n}\mathbf{e}_1\mathbf{n} H_2^3(\mathbf{u}) I_2 \rangle_0}{4} \\ &= \frac{3\,\mathbf{n} \wedge H_2^1(\mathbf{u}) + (\mathbf{n}\mathbf{e}_1\mathbf{n}\mathbf{e}_1\mathbf{n}) \wedge H_2^3(\mathbf{u})}{4} I_2 \end{aligned} \tag{5.57}$$

$$\begin{aligned} W_4(\mathbf{u}) = -W_4'(\mathbf{u}) \cdot \mathbf{e}_2 &= \frac{3\langle \mathbf{n} H_2^1(\mathbf{u}) \rangle_0 - \langle \mathbf{n}\mathbf{e}_1\mathbf{n}\mathbf{e}_1\mathbf{n} H_2^3(\mathbf{u}) \rangle_0}{4} \\ &= \frac{3\,\mathbf{n} \cdot H_2^1(\mathbf{u}) - (\mathbf{n}\mathbf{e}_1\mathbf{n}\mathbf{e}_1\mathbf{n}) \cdot H_2^3(\mathbf{u})}{4} \ . \end{aligned} \tag{5.58}$$

These frequency responses are designed to produce the quadrature components to the windowed signals, which are obtained by the frequency responses W_1 and W_2. However, a fundamental property of quadrature components is violated since

[13] Note that the frequency response of the Riesz transform at the frequency vector $\mathbf{u}$ points in the direction $\mathbf{u}I_2^{-1}$. Since the Riesz transform is steered by the orientation in this case, the relevant information is located on the $\mathbf{e}_1$-axis, and hence it must be projected onto $\mathbf{e}_1 I_2^{-1} = -\mathbf{e}_2$.

the amplitude responses are not identical due to the projections from W_3' to W_3 and from W_4' to W_4. These projections are necessary and since they must be linear[14], they cannot give exactly the same amplitude responses. Nevertheless, the quadrature components can be improved by optimizing the linear factors of the first and third harmonic. As a constraint the sum of the two coefficients must be one, otherwise the quadrature relation would not be exact for signals which obey ideally the 2D signal model. By minimizing the mean square distance between the ideal frequency response and the linear combination of first and third order harmonic, the following solution is obtained:

$$W_3(\mathbf{u}) = \left(\left(\frac{1}{2}+\frac{16}{15\pi}\right)\mathbf{n}\wedge H_2^1(\mathbf{u}) + \left(\frac{1}{2}-\frac{16}{15\pi}\right)(\mathbf{n}\mathbf{e}_1\mathbf{n}\mathbf{e}_1\mathbf{n})\wedge H_2^3(\mathbf{u})\right) I_2 \quad (5.59)$$

$$W_4(\mathbf{u}) = \left(\frac{1}{2}+\frac{16}{15\pi}\right)\mathbf{n}\cdot H_2^1(\mathbf{u}) - \left(\frac{1}{2}-\frac{16}{15\pi}\right)(\mathbf{n}\mathbf{e}_1\mathbf{n}\mathbf{e}_1\mathbf{n})\cdot H_2^3(\mathbf{u}) \ . \quad (5.60)$$

This approximation yields a mean square error of only 0.1% of the mean energy.

In order to close the loop to the discussion about the basis functions $\Psi_1,\ldots,\Psi_4$, i.e., in terms of even and odd symmetries, the frequency responses $W_1,\ldots,W_4$ are combined to obtain these symmetries. The even-even symmetry of $\Psi_1(\mathbf{u})$ is trivial, it is given by a constant function. The odd-odd symmetry frequency response is already known from (5.48) and it is obtained by

$$\Psi_4(\mathbf{u}) = W_1(\mathbf{u}) - W_2(\mathbf{u}) \ .$$

Similar linear combinations of W_3 and W_4 yield the remaining two frequency responses

$$\Psi_2(\mathbf{u}) = W_3(\mathbf{u}) + W_4(\mathbf{u}) \ \text{ and } \ \Psi_3(\mathbf{u}) = W_3(\mathbf{u}) - W_4(\mathbf{u}) \ .$$

The three frequency responses can be found in figure 5.14.

Finally, the whole filter design process is summarized in figure 5.15, where the upper part shows rather the derivation of w_3 and w_4, whereas the bottom part shows the most efficient calculation scheme. These graphs can also be used for implementing the structure multivector.

5.5.2 Amplitude and Phase Decomposition

Up to now, the structure multivector M_S has only been used to extract the orientation information. However, the really interesting features which have to be extracted are the local phases and the local amplitudes of the two i1D partial signals. Based on these four features further properties can be derived, e.g., the local

[14] Note that the steering operations and projections are really performed in spatial domain – the discussion in the frequency domain is only for the sake of simplicity. The design in the frequency domain must be linear in order to transfer it to the spatial domain.

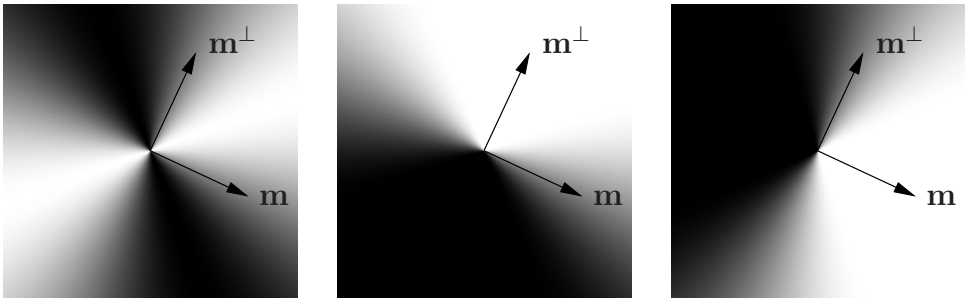

Fig. 5.14: Frequency responses for symmetry decomposition obtained by linear combinations of steered spherical harmonics. From left to right: odd-odd symmetry (Ψ_4), even-odd symmetry (Ψ_2), and odd-even symmetry (Ψ_3).

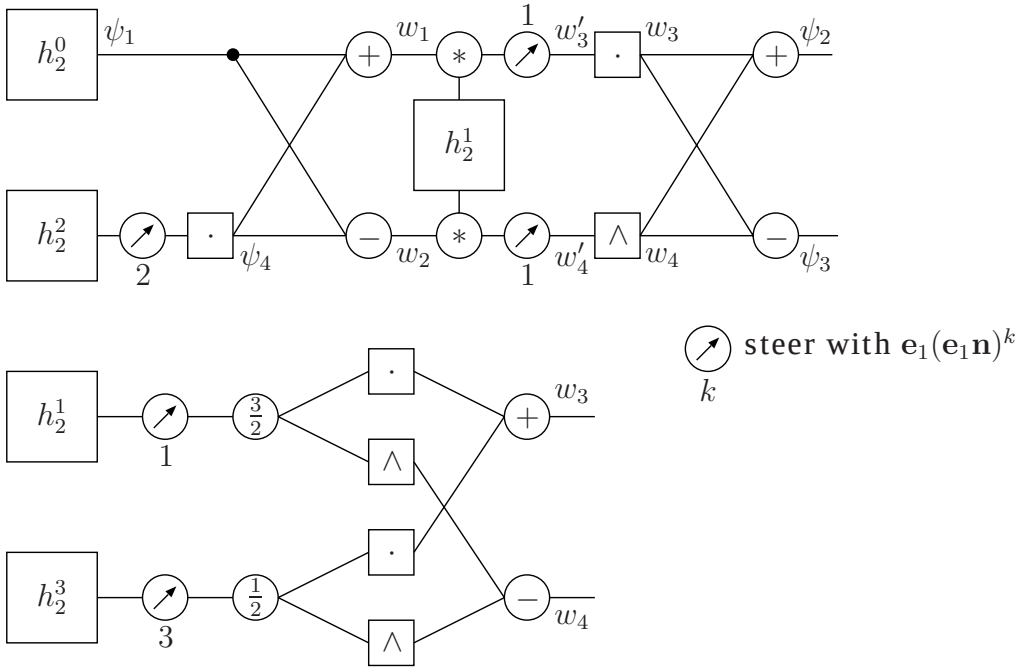

Fig. 5.15: Overview of the filter design process. The upper part shows the derivation of ψ_2 and ψ_3 from the basis filters ψ_1 and ψ_4, whereas the bottom part shows the direct formulation. The linear factors $3/2$ and $1/2$ can be replaced by the optimized factors (see text). The wedge and dot products represent the projections, see also (5.61)–(5.64).

isotropy and the local total amplitude. Obviously, all these features are rotation invariant, since they are extracted from the steered filter outputs.

The local amplitudes of the two i1D partial signals are obtained from the operator outputs of $\mathcal{W}_1, \ldots, \mathcal{W}_4$ which read in terms of M_{S}

$$\begin{aligned}
\mathcal{W}_1\mathbf{f} &= \frac{2M_3 + (\mathbf{n}\mathbf{e}_1)^2(M_0 + M_{12}I_2) + (\mathbf{e}_1\mathbf{n})^2(M_0 - M_{12}I_2)}{4} \\
&= \frac{M_3 + (\mathbf{n}\mathbf{e}_1\mathbf{n}) \cdot (M_0\mathbf{e}_1 + M_{12}\mathbf{e}_2)}{2} = \frac{M_3 + \cos(2\theta_{\mathrm{e}})M_0 + \sin(2\theta_{\mathrm{e}})M_{12}}{2} \qquad (5.61) \\
\mathcal{W}_2\mathbf{f} &= \frac{2M_3 - (\mathbf{n}\mathbf{e}_1)^2(M_0 + M_{12}I_2) - (\mathbf{e}_1\mathbf{n})^2(M_0 - M_{12}I_2)}{4} \\
&= \frac{M_3 - \cos(2\theta_{\mathrm{e}})M_0 - \sin(2\theta_{\mathrm{e}})M_{12}}{2} \qquad (5.62) \\
\mathcal{W}_3\mathbf{f} &= \frac{3\mathbf{n} \cdot (M_1\mathbf{e}_1 + M_2\mathbf{e}_2)I_2 + (\mathbf{n}\mathbf{e}_1\mathbf{n}\mathbf{e}_1\mathbf{n}) \cdot (M_{31}\mathbf{e}_1 - M_{23}\mathbf{e}_2)I_2}{4} \\
&= \frac{3(\cos(\theta_{\mathrm{e}})M_1 + \sin(\theta_{\mathrm{e}})M_2) + \cos(3\theta_{\mathrm{e}})M_{31} - \sin(3\theta_{\mathrm{e}})M_{23}}{4} I_2 \qquad (5.63) \\
\mathcal{W}_4\mathbf{f} &= \frac{3\mathbf{n} \wedge (M_1\mathbf{e}_1 + M_2\mathbf{e}_2) - (\mathbf{n}\mathbf{e}_1\mathbf{n}\mathbf{e}_1\mathbf{n}) \wedge (M_{31}\mathbf{e}_1 - M_{23}\mathbf{e}_2)}{4} \\
&= \frac{3(-\sin(\theta_{\mathrm{e}})M_1 + \cos(\theta_{\mathrm{e}})M_2) + \sin(3\theta_{\mathrm{e}})M_{31} + \cos(3\theta_{\mathrm{e}})M_{23}}{4} I_2 \qquad (5.64)
\end{aligned}$$

Hence, the two local amplitudes are given by

$$\begin{aligned}
A_1(\mathbf{x}) &= |\mathcal{W}_1\mathbf{f}(\mathbf{x}) + \mathcal{W}_3\mathbf{f}(\mathbf{x})| \text{ and} \qquad (5.65) \\
A_2(\mathbf{x}) &= |\mathcal{W}_2\mathbf{f}(\mathbf{x}) + \mathcal{W}_4\mathbf{f}(\mathbf{x})| \ . \qquad (5.66)
\end{aligned}$$

By means of these two amplitudes, it is possible to define the *dominance index* $d(\mathbf{x})$ which is given by the index of the larger amplitude at every position:

$$d(\mathbf{x}) = \begin{cases} 1 & \text{if } A_1(\mathbf{x}) \geq A_2(\mathbf{x}) \\ 2 & \text{else} \end{cases} \ . \qquad (5.67)$$

Using this index, it is possible to change the estimated orientation into the local main orientation, or *major orientation*,

$$\theta_0(\mathbf{x}) = \begin{cases} \theta_{\mathrm{e}}(\mathbf{x}) & \text{if } d(\mathbf{x}) = 1 \\ \theta_{\mathrm{e}}(\mathbf{x}) + \pi/2 & \text{else} \end{cases} \ . \qquad (5.68)$$

One can also select the larger amplitude (i.e., the one which corresponds to the major orientation), the *major amplitude* $A_{\mathrm{M}}(\mathbf{x})$, or the smaller amplitude (i.e., the one which corresponds to the minor orientation), the *minor amplitude* $A_{\mathrm{m}}(\mathbf{x})$, by

$$A_{\mathrm{M}}(\mathbf{x}) = A_{d(\mathbf{x})}(\mathbf{x}) \text{ and } A_{\mathrm{m}}(\mathbf{x}) = A_{3-d(\mathbf{x})}(\mathbf{x}) \ , \qquad (5.69)$$

which is especially useful for designing feature detectors (see section 6.5). The sum of both amplitudes yields the *total amplitude*

$$A_{\mathrm{t}}(\mathbf{x}) = A_{\mathrm{M}}(\mathbf{x}) + A_{\mathrm{m}}(\mathbf{x}) = A_1(\mathbf{x}) + A_2(\mathbf{x}) \ . \tag{5.70}$$

Similar to the discussion about the relationship between the coherence and the isotropy in section 5.2.3, it is worth to look at the relationship between the total amplitude, the determinant and the trace of the structure / metric tensor, and the total scale. To some extent, the two amplitudes A_{M} and A_{m} can be identified with the eigenvalues of the metric tensor (or the structure tensor). Hence, the total amplitude corresponds to the trace of the tensor. On the other hand, the total (isotropic) scale according to (5.16) contains the product of the two ellipse axes, and therefore it contains the product of the eigenvalues, i.e., the determinant of the tensor. Consequently, the total scale could be estimated by the fourth root of the product of the two amplitudes. However, further investigations on this topic are necessary, before a generalized scale adaptive approach can be proposed (see also section 4.6.3).

Same as the total scale, the *local isotropy* can be estimated by the local amplitudes. The estimate is given by the square root[15] of the quotient of the smaller amplitude and the larger amplitude (see also (5.14)):

$$c(\mathbf{x}) = \sqrt{A_{\mathrm{m}}(\mathbf{x})/A_{\mathrm{M}}(\mathbf{x})} \ . \tag{5.71}$$

The features left to be defined are the local i1D phases which read

$$\varphi_1(\mathbf{x}) = \arg(\mathcal{W}_1\mathbf{f}(\mathbf{x}) + \mathcal{W}_3\mathbf{f}(\mathbf{x})) \text{ and} \tag{5.72}$$
$$\varphi_2(\mathbf{x}) = \arg(\mathcal{W}_2\mathbf{f}(\mathbf{x}) + \mathcal{W}_4\mathbf{f}(\mathbf{x})) \ . \tag{5.73}$$

By use of the dominance index, the two phases can be assigned to the *major phase* and the *minor phase*:

$$\varphi_{\mathrm{M}}(\mathbf{x}) = \varphi_{d(\mathbf{x})}(\mathbf{x}) \text{ and} \tag{5.74}$$
$$\varphi_{\mathrm{m}}(\mathbf{x}) = \varphi_{3-d(\mathbf{x})}(\mathbf{x}) \ . \tag{5.75}$$

Note that both phases are obtained in the whole interval $[-\pi, \pi)$ since the two Riesz components obtained by $\mathcal{W}_3'$ and $\mathcal{W}_4'$ are projected onto the upper halfspace. It is assumed that the directions of the two partial signals are equal to their orientations, i.e., they lie in the first and second quadrant. Obviously, the chosen directions may be wrong with respect to intra- or inter-image constraints (see section 4.4) such that one or both directions lie in the third or fourth quadrant. In that case the same orientation-phase wrapping occurs as in the case of the monogenic signal. The necessary corrections are exactly the same as described in section 4.4.

[15] The amplitudes are identified with the eigenvalues of the metric tensor (see above). Therefore, their ratio determines the square of the isotropy, see (5.14).

In figure 5.16 the i2D phase decomposition is illustrated by a synthetic example, which is basically the egg-tray signal, but with different frequencies, different amplitudes, an 'incompatible' orientation (wrt. the coordinate axes), and additive noise.

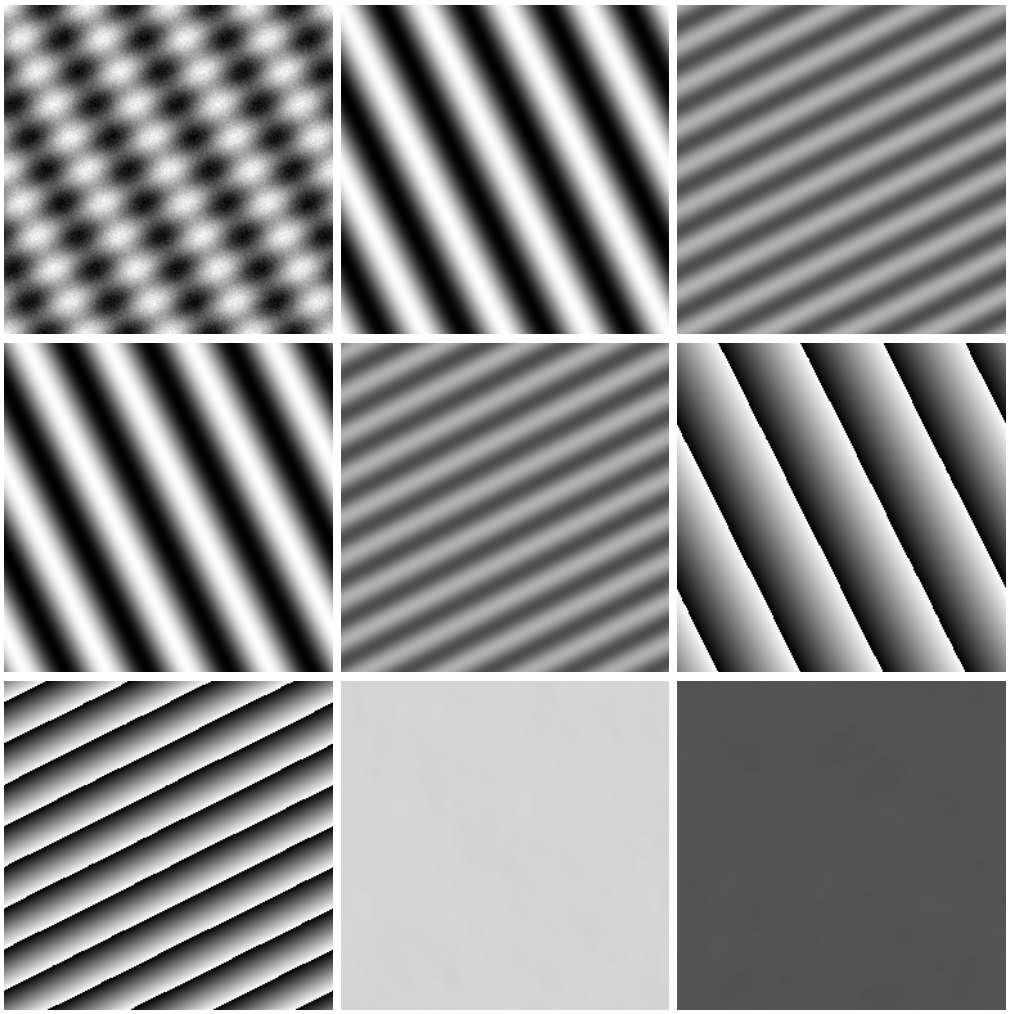

Fig. 5.16: Phase decomposition of noisy image (16.8dB) with different local frequencies and different local amplitudes in orthogonal orientations. Upper row from left to right: test image, operator output of $\mathcal{W}_1$, operator output of $\mathcal{W}_2$. Center row from left to right: operator output of $\mathcal{W}_3$, operator output of $\mathcal{W}_4$, local major phase. Bottom row from left to right: local minor phase, local major amplitude, and local minor amplitude. The local amplitudes are nearly constant, exactly as they should be.

5.5.3 The Properties of the Structure Multivector

In the previous section, a set of 2D features has been extracted from the structure multivector which is much richer than a simple 'amplitude and phase decomposition'. A fundamental property of this feature set is that it fulfills the invariance – equivariance requirement. To show this, the feature set has been computed for another test signal, see figure 5.17. Instead of a Cartesian i2D signal, a log-polar modulation has been chosen (the superposition of two basis functions of the Fourier-Mellin transform [51]). It is quite obvious, that all features are mutually *independent*. A small exception is again the orientation-phase wrapping. At the line where the major orientation wraps from zero to π, the major phase is inverted. The same occurs for the minor orientation (perpendicular to the major orientation) and the minor phase.

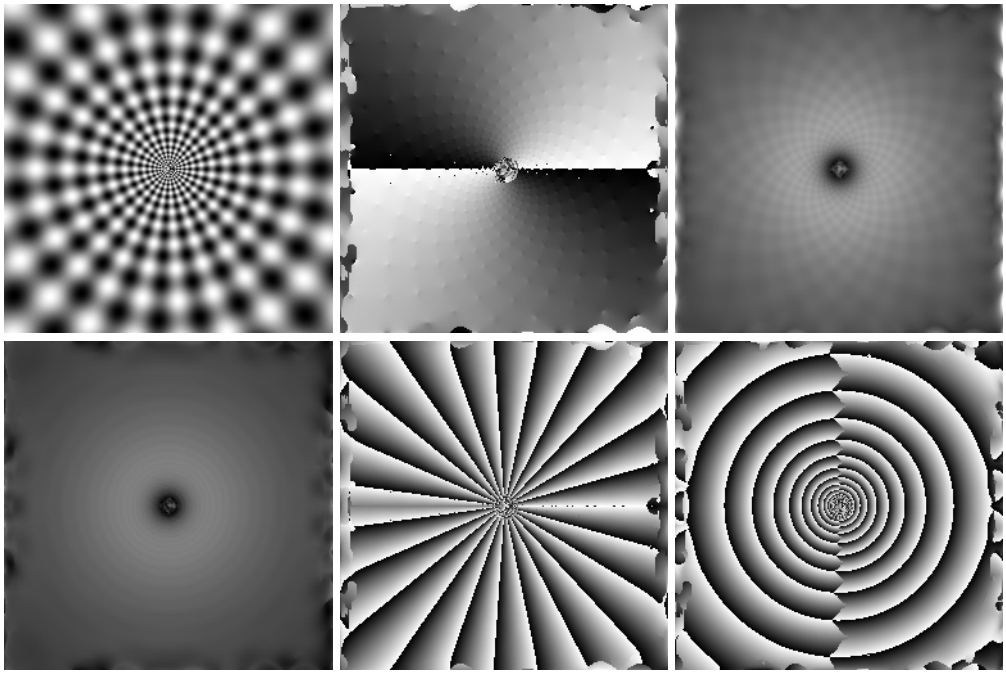

Fig. 5.17: First row, from left to right: second test image (the superposition of two basis functions of the Fourier-Mellin transform), estimated major orientation, and major amplitude. Second row, from left to right: minor amplitude, major phase, and minor phase.

The extracted local features have direct semantic interpretations. Therefore, the further processing based on the SMV output is much simplified, as I show in the applications, chapter 6. The derived features describe the signal by means of symmetry and can be considered as a basis for further semi-local associations. As in the

case of the 1D analytic signal, the semantically relevant information is not directly contained in the filter responses, but in a transformed representation of the output, i.e., in a different (adaptive) coordinate system. Therefore they are compatible with more sophisticated definitions of features, see e.g. [80].

For images which obey the 2D signal model, a *complete* representation is obtained, and hence the SMV performs a split of identity. To be more concrete, the signal is represented by means of five independent features for each scale:

- two local amplitudes
- the local (main) orientation and
- two local i1D phases.

Note that the reduction of dimensionality from the 7D SMV to the five scalar features is obtained by two projections $W_3(\mathbf{u}) = -W_3'(\mathbf{u})\cdot\mathbf{e}_2$ and $W_4(\mathbf{u}) = -W_4'(\mathbf{u})\cdot\mathbf{e}_2$, see (5.57) and (5.58). The remaining rejections might also contain further information.

For many applications it is reasonable to compute also some non-independent features and store them with the SMV:

- the local isotropy
- the local total amplitude
- the local dominance index
- the minor orientation and
- the local scale at which the SMV was computed.

The decomposition according to the local scale depends on the chosen bandpass filters. These filters can either be taken from a filter bank, or one can try to estimate the intrinsic scales. The latter method should also contain anisotropic scales, applying the filters derived in section 5.1–5.3. However, the adaptive choice of anisotropic scale by means of the SMV is still under investigation and will be the topic of future work.

If the considered signal is i1D, the minor part of the SMV is zero. Hence, the main orientation and the total amplitude are given solely by the present i1D signal. In this case the SMV is basically identical to the monogenic signal. In practice no ideal i1D signals exist, but the isotropy can be quite close to zero. The optimal amplitude response in this (nearly) i1D case is highly anisotropic. By means of the isotropy c and the main orientation, the bandpass characteristic can directly be steered according to the approximation in section 5.3.

If the signal is i2D, the isotropy is close to one, and hence the corresponding optimal bandpass is simply given by the isotropic one. However, the full feature set of the SMV is relevant and should be considered in the further processing of the data.

Summarizing these two cases, the following tradeoff between the amplitude response and the phase interpretation is obtained. For i1D signals, the simple phase approach (i.e., one i1D phase) is combined with the more complex (i.e., anisotropic) amplitude response whereas for i2D signals, the more complex phase approach (i.e., full SMV information) is combined with the simpler (i.e., isotropic) amplitude response (see figure 5.18).

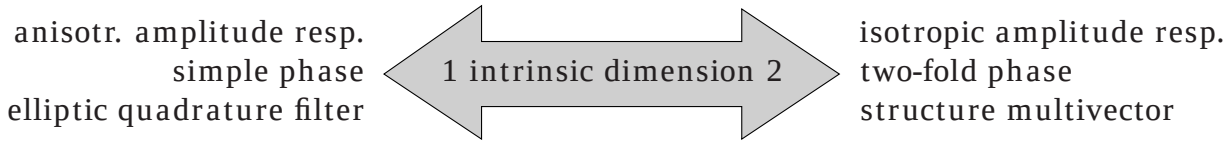

Fig. 5.18: Tradeoff between complex phase approach / simple amplitude response and simple phase approach / complex amplitude response.

The SMV combines ideas and properties from various approaches of low-level signal processing, as there are the structure tensor, the (quaternionic) analytic signal, and scale-space, into one single and coherent filter design principle. The full feature set obtained from the structure tensor is contained in the SMV. Furthermore, the quaternionic analytic signal drops out as a special case of the SMV with fixed orientation, although the phase model is different, since the quaternionic phase approach is based on the functions $\Psi_1, \ldots, \Psi_4$ instead of $W_1, \ldots, W_4$. However, with respect to robustness the latter choice is preferable, since violations of the assumed signal model are smoothly windowed by the $\cos^2$-terms instead of having rectangular profiles.

5.6 Summary of Chapter 5

The tradeoff mentioned above is the final outcome of this chapter, in which the Laplace equation has been extended to adaptive local metrics and in which a phase approach for i2D signals has been developed.

The extension to local metrics leads to an anisotropic bandpass filter design which has a linear computational complexity since the filters can be approximated and steered. The formulas which are necessary for computing these elliptic quadrature filters are summarized in table 5.2. The local metric also leads to the concept of local isotropy which replaces the feature coherence in this framework.

basis filter responses in the Fourier domain
$G_0^0(\mathbf{u}) = \exp(-2\pi\lvert\mathbf{u}\rvert s')\mathbf{F}(\mathbf{u})$
$G_1^0(\mathbf{u}) = \frac{\mathbf{u}}{\lvert\mathbf{u}\rvert} I_2^{-1} \exp(-2\pi\lvert\mathbf{u}\rvert s')\mathbf{F}(\mathbf{u})$
$G_1^1(\mathbf{u}) = \mathbf{u}\, I_2^{-1} \exp(-2\pi\lvert\mathbf{u}\rvert s')\mathbf{F}(\mathbf{u})$
$G_2^1(\mathbf{u}) = \frac{\mathbf{u}\mathbf{e}_1\mathbf{u}}{\lvert\mathbf{u}\rvert} \exp(-2\pi\lvert\mathbf{u}\rvert s')\mathbf{F}(\mathbf{u})$
$G_3^0(\mathbf{u}) = \frac{\mathbf{u}\mathbf{e}_1\mathbf{u}\mathbf{e}_1\mathbf{u}}{\lvert\mathbf{u}\rvert^3} I_2^{-1} \exp(-2\pi\lvert\mathbf{u}\rvert s')\mathbf{F}(\mathbf{u})$
$G_3^1(\mathbf{u}) = \frac{\mathbf{u}\mathbf{e}_1\mathbf{u}\mathbf{e}_1\mathbf{u}}{\lvert\mathbf{u}\rvert^2} I_2^{-1} \exp(-2\pi\lvert\mathbf{u}\rvert s')\mathbf{F}(\mathbf{u})$
approximations of anisotropic filter responses
$(l(c) * \mathbf{f})(\mathbf{x}) \approx g_0^0(\mathbf{x}) + (c-1)\pi s' \,(\mathbf{n}\mathbf{e}_1\mathbf{n}) \cdot g_2^1(\mathbf{x})$
$(h(c) * l(c) * \mathbf{f})(\mathbf{x}) \approx g_1^0(\mathbf{x}) + (c-1)\mathbf{e}_1 \left(\frac{1}{4}(\mathbf{n}\mathbf{e}_1\mathbf{n} g_3^0(\mathbf{x}) - g_1^0(\mathbf{x})\mathbf{n}\mathbf{e}_1\mathbf{n}) + \frac{\pi s'}{2}(g_1^1(\mathbf{x})\mathbf{n}\mathbf{e}_1\mathbf{n} + \mathbf{n}\mathbf{e}_1\mathbf{n} g_3^1(\mathbf{x}))\right)$

Tab. 5.2: Formulas for computing the elliptic quadrature filters.

The i2D phase approach leads to a method for estimating the local metric and for describing the local structure of i2D signals. Based on a 2D signal model the structure multivector is used to extract a rich feature set, consisting of major, minor, and total amplitude, isotropy, major and minor phase, and main (major) orientation. The features are based on the output of certain filters, which can be found in table 5.3.

basis filter responses ($H_2^k(\mathbf{u})$ are the harmonics of order k)

$$M_{\mathrm{S}}(\mathbf{x}) = \mathbf{f}(\mathbf{x}) + (h_2^1 * \mathbf{f})(\mathbf{x}) + \mathbf{e}_3(h_2^2 * \mathbf{f})(\mathbf{x}) + \mathbf{e}_3(h_2^3 * \mathbf{f})(\mathbf{x})$$

$$\mathcal{W}_1\mathbf{f}(\mathbf{x}) = \tfrac{1}{2}(M_3(\mathbf{x}) + \cos(2\theta_{\mathrm{e}})M_0(\mathbf{x}) + \sin(2\theta_{\mathrm{e}})M_{12}(\mathbf{x}))$$

$$\mathcal{W}_2\mathbf{f}(\mathbf{x}) = \tfrac{1}{2}(M_3(\mathbf{x}) - \cos(2\theta_{\mathrm{e}})M_0(\mathbf{x}) - \sin(2\theta_{\mathrm{e}})M_{12}(\mathbf{x}))$$

$$\mathcal{W}_3\mathbf{f}(\mathbf{x}) = \tfrac{1}{4}(3(\cos(\theta_{\mathrm{e}})M_1(\mathbf{x}) + \sin(\theta_{\mathrm{e}})M_2(\mathbf{x})) + \cos(3\theta_{\mathrm{e}})M_{31}(\mathbf{x}) - \sin(3\theta_{\mathrm{e}})M_{23}(\mathbf{x}))I_2$$

$$\mathcal{W}_4\mathbf{f}(\mathbf{x}) = \tfrac{1}{4}(3(-\sin(\theta_{\mathrm{e}})M_1(\mathbf{x}) + \cos(\theta_{\mathrm{e}})M_2(\mathbf{x})) + \sin(3\theta_{\mathrm{e}})M_{31}(\mathbf{x}) + \cos(3\theta_{\mathrm{e}})M_{23}(\mathbf{x}))I_2$$

features (note that θ_{e} is estimated before $\mathcal{W}_k\mathbf{f}$ is evaluated)

$$\theta_{\mathrm{e}}(\mathbf{x}) = \tfrac{1}{4}\arg((M_0(\mathbf{x}) + M_{12}(\mathbf{x})I_2)^2 + (M_{31}(\mathbf{x}) - M_{23}(\mathbf{x})I_2)(M_1(\mathbf{x}) + M_2(\mathbf{x})I_2))$$

$$A_1(\mathbf{x}) = |\mathcal{W}_1\mathbf{f}(\mathbf{x}) + \mathcal{W}_3\mathbf{f}(\mathbf{x})| \quad \text{and} \quad A_2(\mathbf{x}) = |\mathcal{W}_2\mathbf{f}(\mathbf{x}) + \mathcal{W}_4\mathbf{f}(\mathbf{x})|$$

$$\varphi_1(\mathbf{x}) = \arg(\mathcal{W}_1\mathbf{f}(\mathbf{x}) + \mathcal{W}_3\mathbf{f}(\mathbf{x})) \quad \text{and} \quad \varphi_2(\mathbf{x}) = \arg(\mathcal{W}_2\mathbf{f}(\mathbf{x}) + \mathcal{W}_4\mathbf{f}(\mathbf{x}))$$

$$d(\mathbf{x}) = \begin{cases} 1 & \text{if } A_1(\mathbf{x}) \geq A_2(\mathbf{x}) \\ 2 & \text{else} \end{cases} \quad \text{and} \quad \theta_0(\mathbf{x}) = \begin{cases} \theta_{\mathrm{e}}(\mathbf{x}) & \text{if } d(\mathbf{x}) = 1 \\ \theta_{\mathrm{e}}(\mathbf{x}) + \pi/2 & \text{else} \end{cases}$$

$$A_{\mathrm{M}}(\mathbf{x}) = A_{d(\mathbf{x})}(\mathbf{x}) \quad \text{and} \quad A_{\mathrm{m}}(\mathbf{x}) = A_{3-d(\mathbf{x})}(\mathbf{x})$$

$$A_{\mathrm{t}}(\mathbf{x}) = A_{\mathrm{M}}(\mathbf{x}) + A_{\mathrm{m}}(\mathbf{x}) \quad \text{and} \quad c(\mathbf{x}) = \sqrt{A_{\mathrm{m}}(\mathbf{x})/A_{\mathrm{M}}(\mathbf{x})}$$

$$\varphi_{\mathrm{M}}(\mathbf{x}) = \varphi_{d(\mathbf{x})}(\mathbf{x}) \quad \text{and} \quad \varphi_{\mathrm{m}}(\mathbf{x}) = \varphi_{3-d(\mathbf{x})}(\mathbf{x})$$

Tab. 5.3: Formulas for computing the features of the SMV.

Chapter 6

APPLICATIONS

'How can it be that mathematics,
being after all a product of human thought
independent of experience,
is so admirably adapted to the objects of reality?'

Albert Einstein (1879-1955)

In this chapter several applications of the derived framework of low level image processing are presented. Although the applications are not the central point of this rather theoretic thesis, this chapter underlines the applicability of the developed theory. The derived methods cover quite different fields as orientation estimation, line and edge detection, stereo correspondence and disparity estimation, adaptive smoothing, and corner detection. The comparatively good performance of the presented methods proves the capability of the fundamental theoretic framework. All applications have in common that they use only few parameters and that they are obtained by a straightforward derivation from the theoretic considerations in the previous chapters.

6.1 Orientation Estimation

Basically, there are two ways to estimate the local orientation by means of the proposed signal preprocessing methods. The first approach is based on the monogenic signal solely, which means that only the (bandpass filtered) signal and its Riesz transform are used. The second approach is based on the basis functions of the structure multivector (SMV), i.e., the spherical harmonics of order zero to three. Whereas in the first case an averaging scheme must be applied, the orientation can be estimated pointwise in case of the SMV.

Throughout the considerations about orientation estimation, it is assumed that the images contain additive noise. Although the model of additive noise is quite

incomplete, it covers many other noise characteristics if the signal-to-noise ratio (SNR) is varied locally, e.g. multiplicative noise is covered by this method, at least approximately. For the experiments I solely use synthetic signals, since for real images the ground truth and the noise level are unknown. The used test signal is a radial modulation (see figure 6.1) with added Gaussian distributed white noise.

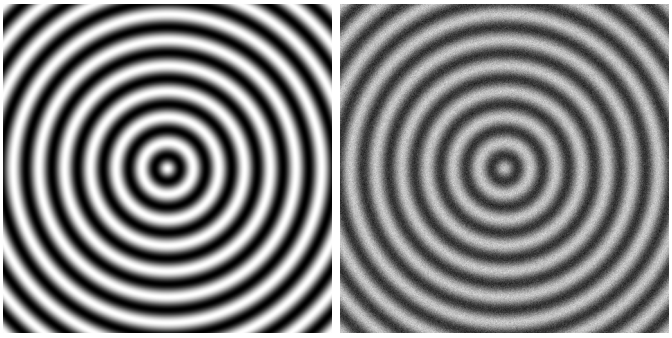

Fig. 6.1: Synthetic test images with known ground truth for optimizing the orientation estimation. Image without noise (left) and with additive noise (SNR 11dB, right).

In the subsequent section, some general considerations about the optimal averaging of orientation information are made. The results are necessary for the orientation estimation by means of the monogenic signal.

6.1.1 Optimal Local Orientation Estimation

Several approaches for orientation estimation are based on gradient operators (e.g. Sobel operator, derivative of Gaussian). These operators do not provide complete orientation maps since even structures (e.g. lines) yield a zero output. The same is true for the Riesz transform, which is the derivative of a harmonic potential. However, the local phase of the monogenic signal is zero or π in the case of an even structure, such that those positions can be identified. The aim of this section and the following one is to derive a *phase adaptive averaging* of the orientation information in order to obtain complete orientation maps. Thus, an optimal local orientation estimator based on averaged responses of a SQF is developed in this section.

The local averaging is based on three local properties: the distance to the origin of the neighborhood, the local amplitude, and the local phase. The dependence on the local distance is realized by some window function, the local amplitude and the local phase are directly obtained from the SQF response. The local amplitude can be considered as a measure for the local intensity. A high local intensity means reliable local features. On the other hand, the local phase realizes a measure for the

local uncertainty of the orientation estimation, i.e., it represents the ratio between odd and even structures. Hence, cases occur where the orientation estimation is not reliable although the intensity is high. Combining the three local properties yields the following general weighting function:

$$g(\mathbf{x}-\mathbf{x}_0) = w(\mathbf{x}-\mathbf{x}_0)m(A(\mathbf{x}-\mathbf{x}_0))p(\varphi(\mathbf{x}-\mathbf{x}_0)) \ , \tag{6.1}$$

where w indicates the used window, m indicates a function of the local amplitude A, and p indicates a function of the local phase φ. Applying this weighting function yields the general optimization problem

$$\min \sum_{\mathbf{x}\in\mathcal{N}_{\mathbf{x}_0}} g(\mathbf{x}-\mathbf{x}_0)((\mathbf{n}_0\cdot\mathbf{n}_\mathbf{x})^2 - \mathbf{n}_\mathbf{x}^2) \ , \tag{6.2}$$

where $\mathcal{N}_{\mathbf{x}_0}$ is a local neighborhood of $\mathbf{x}_0$, $\mathbf{n}_0 = \cos\theta_0\mathbf{e}_1 + \sin\theta_0\mathbf{e}_2$ is the orientation to be optimized, and $\mathbf{n}_\mathbf{x}$ is the unit orientation vector obtained from the SQF at the position $\mathbf{x}$. The optimizer is formulated in such a way that the variance of the orthogonal distance between the estimated vector and the set of orientation vectors is minimized (see figure 6.2).

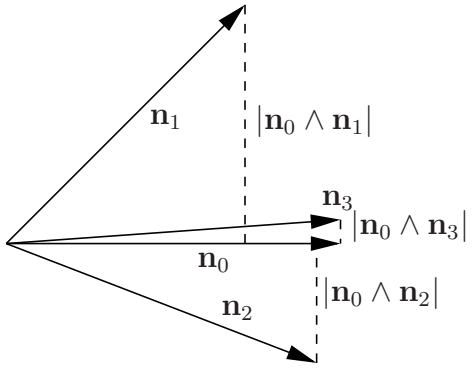

Fig. 6.2: Geometric interpretation of the minimizer. Three unit orientation vectors are given by $\mathbf{n}_1$, $\mathbf{n}_2$, and $\mathbf{n}_3$. The vector $\mathbf{n}_0$ is the estimated orientation. The optimization criterion is the square of the orthogonal distance which is given by $|\mathbf{n}_0 \wedge \mathbf{n}_k|^2$, $k = 1, 2, \ldots$.

Since

$$(\mathbf{n}_0\cdot\mathbf{n}_\mathbf{x})^2 = \frac{1}{4}(\mathbf{n}_0\mathbf{n}_\mathbf{x} + \mathbf{n}_\mathbf{x}\mathbf{n}_0)^2 = \frac{1}{2}\mathbf{n}_\mathbf{x}^2 + \frac{1}{4}(\mathbf{n}_\mathbf{x}\mathbf{n}_0\mathbf{n}_\mathbf{x}\mathbf{n}_0 + \mathbf{n}_0\mathbf{n}_\mathbf{x}\mathbf{n}_0\mathbf{n}_\mathbf{x})$$

the optimization problem can be rewritten as

$$\min \sum_{\mathbf{x}\in\mathcal{N}_{\mathbf{x}_0}} g(\mathbf{x}-\mathbf{x}_0)\left(\frac{1}{2}(\mathbf{n}_\mathbf{x}\mathbf{n}_0\mathbf{n}_\mathbf{x}\mathbf{n}_0 + \mathbf{n}_0\mathbf{n}_\mathbf{x}\mathbf{n}_0\mathbf{n}_\mathbf{x}) - \mathbf{n}_\mathbf{x}^2\right)$$

and thus setting the θ_0-derivative to zero gives (note that $\frac{\partial \mathbf{n}_0}{\partial \theta_0} = \mathbf{n}_0 I_2$):

$$\sum_{\mathbf{x}\in\mathcal{N}_{\mathbf{x}_0}} g(\mathbf{x}-\mathbf{x}_0)(\mathbf{n}_\mathbf{x}\mathbf{n}_0\mathbf{n}_\mathbf{x}\mathbf{n}_0 - \mathbf{n}_0\mathbf{n}_\mathbf{x}\mathbf{n}_0\mathbf{n}_\mathbf{x}) = 0 \ . \tag{6.3}$$

Introducing $\mathbf{e}_1\mathbf{e}_1 = 1$ in appropriate places yields

$$\begin{aligned} 0 &= \sum_{\mathbf{x}\in\mathcal{N}_{\mathbf{x}_0}} g(\mathbf{x}-\mathbf{x}_0)(\mathbf{n}_\mathbf{x}\mathbf{e}_1\mathbf{n}_\mathbf{x}\mathbf{n}_0\mathbf{e}_1\mathbf{n}_0 - \mathbf{n}_0\mathbf{e}_1\mathbf{n}_0\mathbf{n}_\mathbf{x}\mathbf{e}_1\mathbf{n}_\mathbf{x}) \\ &= X\mathbf{n}_0\mathbf{e}_1\mathbf{n}_0 - \mathbf{n}_0\mathbf{e}_1\mathbf{n}_0 X = 2X \wedge \mathbf{n}_0\mathbf{e}_1\mathbf{n}_0 \ , \end{aligned}$$

where $X = \sum_{\mathbf{x}\in\mathcal{N}_{\mathbf{x}_0}} g(\mathbf{x}-\mathbf{x}_0)\mathbf{n}_\mathbf{x}\mathbf{e}_1\mathbf{n}_\mathbf{x}$. The last expression is clearly fulfilled if

$$\mathbf{n}_0\mathbf{e}_1\mathbf{n}_0 = X = \sum_{\mathbf{x}\in\mathcal{N}_{\mathbf{x}_0}} g(\mathbf{x}-\mathbf{x}_0)\mathbf{n}_\mathbf{x}\mathbf{e}_1\mathbf{n}_\mathbf{x} \ . \tag{6.4}$$

The previous derivation is basically the same as in [8], but it is much easier because there is no need to switch between vector notation and complex numbers.

From (6.4) the averaged orientation is obtained by the generalized logarithm:

$$\theta_0 = \langle \log(\mathbf{n}_0\mathbf{e}_1\mathbf{n}_0\mathbf{e}_1)\rangle_2/2 = \langle \log(X\mathbf{e}_1)\rangle_2/2 \ .$$

Depending on the choice of the weighting function, the final expression can be very simple. In [44], for example, the window function was given as a 3×3 boxfilter, the magnitude was not taken into account, and the phase adaptivity was given by $p(\varphi) = \sin^2\varphi$. As I will show in the subsequent section, this is the optimal choice for the case of noiseless signals. If the magnitude is additionally taken into account with its square, (6.4) is identical to the orientation vector obtained from the structure tensor (if the Sobel operator is replaced by the difference of conjugate Poisson filters, see section 4.6.2)[1]:

$$\mathbf{n}_0\mathbf{e}_1\mathbf{n}_0 = \sum_{\mathbf{x}\in\mathcal{N}_{\mathbf{x}_0}} w(\mathbf{x}-\mathbf{x}_0)(c*\mathbf{f})(\mathbf{x})\mathbf{e}_1(c*\mathbf{f})(\mathbf{x}) \ .$$

6.1.2 Orientation Estimation by the Riesz Transform

As stated above, the orientation information of the monogenic signal must be averaged in order to interpolate the orientation at those positions where the estimate is not reliable. In [18] a similar investigation is presented where the author proposes

[1] According to [126, 75] a filter with sine-shaped radial frequency response is optimal for orientation estimation with the structure tensor. For an appropriate choice of parameters a DOCP filter can have a quite similar shape.

to use $|\sin\varphi|^{0.1183}$ to weight the interpolation. This result is obtained by optimizing the averaging for a set of images with known ground truth. The images he uses are synthetic interferograms (i.e., constant local energy) with constant additive uniform noise (SNR of 15.74dB). Due to this specialized setting, the results cannot be applied for general images.

In contrast to the mentioned experiment above, I vary the SNR in a range from 23dB to -3dB. As it can be seen in figure 6.3, the orientation error depends on the local phase *and* the SNR. The graph shows the logarithm of the mean square ori-

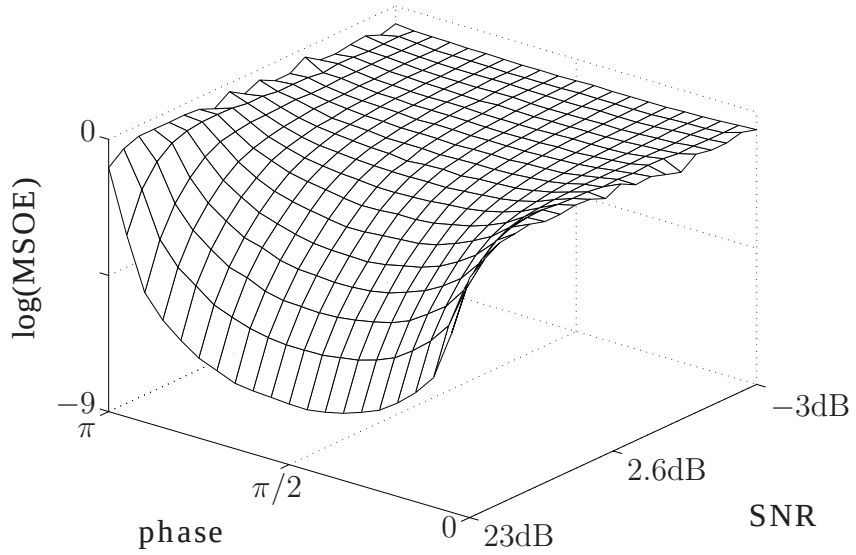

Fig. 6.3: Mean square orientation error (MSOE) in dependence of the local phase and the SNR for additive noise. A value of 0 means an error of $\pi/2$, which is the maximal possible orientation error.

entation error in dependence of the phase angle and the SNR. Actually, the error follows a $|\sin\varphi|^{-k}$-shape. The exponent can be estimated in logarithmic scale by a linear approximation. In table 6.1 the different noise levels and the corresponding optimal exponents k_{e} are listed. Since the approximation is performed with respect to the mean *square* error, the exponents in table 6.1 can directly be used to obtain the function $p(\varphi)$ in (6.1). The weighting function g is multiplied with the double angle orientation vector (see (6.2)), and hence the reliability of this 'square' of the orientation vectors is anti-proportional to their mean square error (this is a special case of the Mahalanobis metric [9]). The exponents are obtained by solving the linear equation

$$k\log(|\sin\varphi|^{-1}) + m = \log(\mathrm{MSOE}) \tag{6.5}$$

where MSOE indicates the mean square orientation error and m indicates some additive offset. Using the exponents from the solutions of the linear systems, an

SNR	23	17	13	11	9.0	7.5	6.1	5.0	3.9	3.0
k_{e}	1.90	1.68	1.49	1.31	1.15	0.97	0.81	0.69	0.59	0.49
k_{a}	1.92	1.69	1.47	1.28	1.11	0.96	0.83	0.71	0.61	0.51
SNR	2.2	1.4	0.7	0.1	-0.5	-1.1	-1.6	-2.1	-2.6	-3.0
k_{e}	0.42	0.37	0.34	0.28	0.24	0.22	0.21	0.18	0.18	0.14
k_{a}	0.44	0.37	0.32	0.28	0.24	0.21	0.19	0.18	0.17	0.16

Tab. 6.1: Relationship between SNR (in dB) and the exponent of $|\sin\varphi|$. The exponents k_{e} are direct estimates from the test images while k_{a} are the exponents obtained from the approximation (6.6).

approximation of the MSOE distribution is obtained (see figure 6.4).

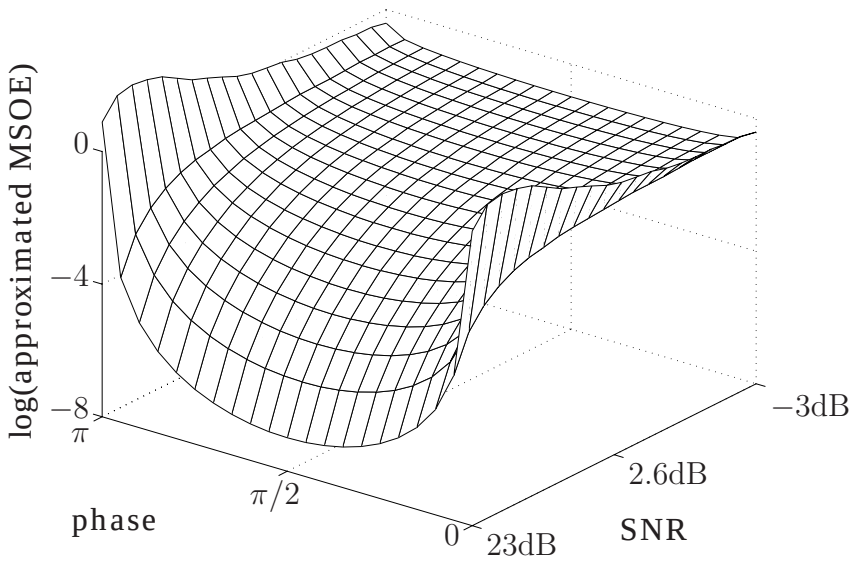

Fig. 6.4: Approximation of the MSOE distribution resulting from the optimal orientation estimation experiment. A value of 0 means an error of $\pi/2$, which is the maximal possible orientation error.

From the measurements k_{e} in table 6.1 the following approximation of the relationship between SNR and k is obtained by a third-order optimization in the log-domain:

$$\begin{aligned} k \approx & -1.8\cdot 10^{-4} e^{-3(\mathrm{SNR}-23.0)/8.69} + 1.2\cdot 10^{-2} e^{-2(\mathrm{SNR}-23.0)/8.69} - 0.27 e^{-(\mathrm{SNR}-23.0)/8.69} + 2.18 \\ = & -0.52\cdot 0.71^{\mathrm{SNR}} + 2.42\cdot 0.79^{\mathrm{SNR}} - 3.80\cdot 0.89^{\mathrm{SNR}} + 2.18 \stackrel{\mathrm{def}}{=} k_{\mathrm{a}} \end{aligned} \tag{6.6}$$

where the SNR is given in dB. This approximation is illustrated in figure 6.5. Obvi-

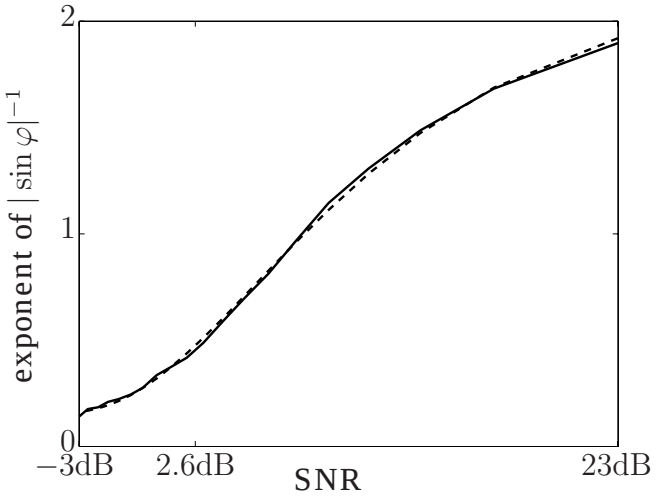

Fig. 6.5: Relation of the exponent k of $|\sin\varphi|^{-1}$ and the SNR. The measurements k_e yield the solid curve and the approximation is indicated by the dashed line.

ously, (6.6) describes the relation of the SNR and the exponent k accurately over a wide range of the SNR. Quite interesting is the asymptotic behavior. For an infinite SNR the optimal exponent is obtained as 2.18 which is pretty close to the theoretic result of 2. This value for k is obtained from the Jacobian of the coordinate transform to polar coordinates which says that a small orientation change $\Delta\theta$ corresponds to an arclength of $R\Delta\theta|\sin\varphi|$ if R is the radius of the sphere. Taking the square difference of θ, the square arclength must be taken into account, which is given by $R^2(\Delta\theta)^2\sin^2\varphi$.

Finally, the orientation variance is depending on the phase and the SNR and it is estimated according to (6.5) as

$$(\Delta\theta)^2 = \exp(k\log(|\sin\varphi|^{-1}) + m) = \exp(m)|\sin\varphi|^{-k} \quad . \tag{6.7}$$

Repeating the same experiment for test images with different radial frequencies gives basically the same results. The exponent k is independent of the frequency in a wide range. The reliability of the orientation information is anti-proportional to the variance and hence proportional to $|\sin\varphi|^k$. If the SNR is locally constant, the factor $\exp(m)$ is not relevant (it cancels out in the averaging). For the application of an optimal orientation estimation, the weighting function should be chosen by (6.6) and for a large range of the SNR the optimal value of k is quite close to two.

For the case cited in the beginning of this section (15.7dB), the optimal averaging is obtained by $|\sin\varphi|^{1.63}$. The discrepancy between the two results can be explained as follows. At first, I apply the optimization to the monogenic signal (allpass) whereas Bülow uses SQFs according to [43] (bandpass). Furthermore, he defines the accuracy measure $\mathrm{acc}_\theta = 1-2\Delta\theta/\pi$ (where $\Delta\theta$ is the difference between the

true orientation and the estimated one) whereas I use the reciprocal standard deviation. The third difference is that I minimize using the L^2-norm, whereas Bülow uses the L^1-norm. Using bandpass filters changes the behavior of the orientation estimation. Basically, one can say that for bandpasses with low center frequency (as it is the case in [18]) the results tend to be better for lower exponents k. The choice of the error measure however, is much more crucial to the optimization. Whereas for allpass filters the two norms basically yield the same results, they differ totally for bandpass filters. The average absolute error is decreased by 30% using the exponent according to [18], but the mean square error is *increased* by nearly 60%. Optimizations in L^2 tend to have less outliers than L^1 optimizations, which is especially useful if some further post processing depends on the orientation information.

6.1.3 Orientation Estimation with Spherical Harmonics

For the structure multivector the spherical harmonics of order zero to three are applied as basis filters, see section 5.4.4. If all these operator outputs are available, it is reasonable to combine them to estimate the orientation. In order to find an appropriate combination, the relationship between the estimation error and the phase must be identified for different signal-to-noise ratios. Therefore, the experiment for the Riesz transform (first order spherical harmonic) from the previous section is repeated for the second and third order spherical harmonic, see figure 6.6 (and compare it to figure 6.3).

Obviously, the second order spherical harmonic shows the opposite behavior of the odd order spherical harmonics. Whereas the latter show an increase of the error for phases close to zero and close to π, the former shows an increase of the error for phases close to $\pi/2$ and $3\pi/2$. Hence, it seems to be reasonable to combine all three operator outputs, in order to obtain an orientation estimate which is independent of the local phase. This is obtained by (5.56) which is basically the sum of the squared second order harmonic and the product of the two odd order harmonics. The resulting MSOE diagram can be found in figure 6.7.

By the following calculation it is verified that (5.56) actually yields the local orientation in the first quadrant (i.e., $\theta_e \in [0, \pi/2)$) independent of the local phase. Consider the signal $\mathbf{f}(\mathbf{x}) = \cos(\mathbf{n} \cdot \mathbf{x})\mathbf{e}_3$ with $\mathbf{e}_1\mathbf{n} = \exp(I_2\theta)$. Its Riesz transform is given by $(h_2^1 * \mathbf{f})(\mathbf{x}) = -\mathbf{n}\sin(\mathbf{n} \cdot \mathbf{x})$. The second order spherical harmonic yields the response $(\mathbf{h}_2^2 * \mathbf{f})(\mathbf{x}) = (\mathbf{e}_1\mathbf{n})^2 \cos(\mathbf{n} \cdot \mathbf{x})\mathbf{e}_3$ and the third order spherical harmonic gives $(\mathbf{h}_2^3 * \mathbf{f})(\mathbf{x}) = -\mathbf{e}_1(\mathbf{e}_1\mathbf{n})^3 \sin(\mathbf{n} \cdot \mathbf{x})$. Hence, the structure multivector reads

$$M_S(\mathbf{x}) = \cos(\mathbf{n} \cdot \mathbf{x})(\mathbf{e}_3 + (\mathbf{e}_1\mathbf{n})^2) - \sin(\mathbf{n} \cdot \mathbf{x})(\mathbf{n} + \mathbf{e}_{31}(\mathbf{e}_1\mathbf{n})^3) \ ,$$

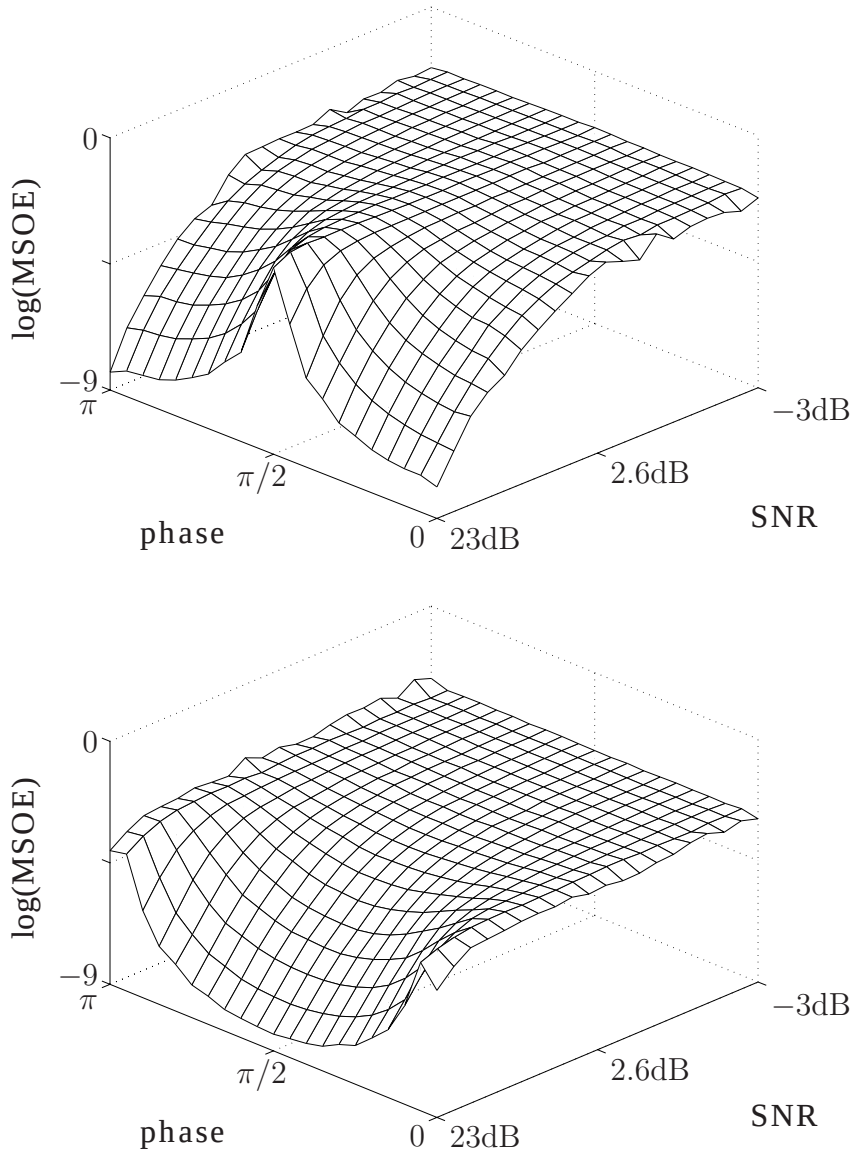

Fig. 6.6: Mean square orientation error in dependence of the local phase and the SNR for additive noise. Top: error of estimation by second order spherical harmonic, bottom: error of estimation by third order spherical harmonic. A value of 0 means an error of $\pi/2$, which is the maximal possible orientation error. Actually, the orientation errors have been evaluated by comparing the outputs of the second and third order harmonic with $(\mathbf{e}_1\mathbf{n})^2$ and $(\mathbf{e}_1\mathbf{n})^3$, yielding a maximal error of $\pi/4$ and $\pi/6$, respectively. Therefore, the error for low SNRs seems to decrease with the order of the harmonic which is just an effect of the measurement.

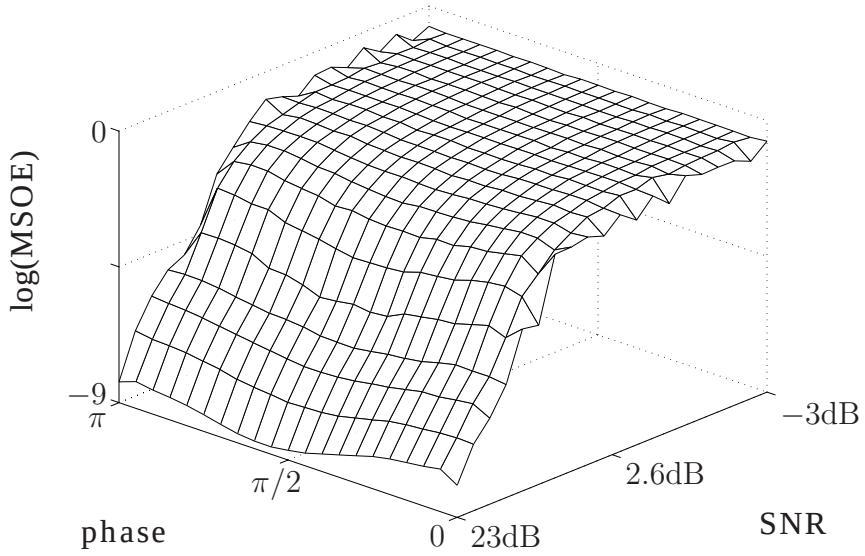

Fig. 6.7: Mean square orientation error of θ_{e} according to (5.56) in dependence of the local phase and the SNR for additive noise. A value of 0 means an error of $\pi/2$, which is the maximal possible orientation error.

and plugging M_{S} into (5.56) results in

$$\begin{aligned}\theta_{\mathrm{e}} &= \frac{1}{4}\arg((M_0 + M_{12}I_2)^2 + (M_{31} - M_{23}I_2)(M_1 + M_2 I_2)) \\ &= \frac{1}{4}\arg(\cos^2(\mathbf{n}\cdot\mathbf{x})(\mathbf{e}_1\mathbf{n})^4 + \sin^2(\mathbf{n}\cdot\mathbf{x})(\mathbf{e}_1\mathbf{n})^3(\mathbf{e}_1\mathbf{n})) \\ &= \frac{1}{4}\langle\log((\mathbf{e}_1\mathbf{n})^4)\rangle_2 = (\theta \mod \pi/2)I_2 \ .\end{aligned}$$

Hence, the estimation by (5.56) is independent of the local phase and yields the orientation in the interval $[0, \pi/2)$.

The orientation estimation according to (5.56) is preferable to the one which is based on the Riesz transform solely, since there is no need to average the estimated orientation due to phase dependences. The only reason to use the estimate by the Riesz transform is to avoid the computation of the higher order filter responses.

6.2 Edge Detection

Standard approaches of edge detection (e.g. the Sobel operator or the Canny detector [28]) apply a certain filter to an image and apply a threshold to the output. The advantage of this method is that for a sufficient signal-to-noise ratio the detection is stable and robust against noise. Unfortunately, the filter output is proportional to the local contrast, and therefore those edges in the image which have less energy than the estimated noise response are eliminated. The reason for this behavior is that the signal with noise is convolved with a certain filter mask which is mostly some kind of derivative operator. These operators are quite sensitive to noise. Hence, a high threshold has to be chosen in order to suppress the noise, with the consequence that the weak edges are not detected. Ordinary thresholds are clearly suboptimal since they do not take into account the *spectral characteristic* of the structure, i.e., if the energy is concentrated on a line. The standard method for improving the performance of the edge detector is to enhance the threshold, e.g. by applying local adaptive thresholds or hysteresis thresholding.

6.2.1 Phase Congruency

Much better results should be obtained if the signal is correlated to a changed version of itself. Since noise is mostly uncorrelated, edges and noise can probably better be distinguished in this way. The question is, however, how this change should look like and how to apply this in an efficient way. One possible group action under which a signal containing an edge is invariant is clearly a shift parallel to the edge. This results in an autocorrelation analysis of the local structure which is equivalent to the structure tensor [54]. Another possible group action is a dilation. Obviously, a rescaled edge is still an edge and if the center of scaling is located on the edge, the scaled versions of the edge are similar. Depending on the noise distribution, both methods reduce the noise contribution better than derivative-based operators. However, it is unclear how to realize an autocorrelation with respect to a scale parameter efficiently.

Considering the frequency domain, the scaling of a signal results in a scaling of its Fourier transform. Hence, being invariant under scaling operations in the spatial domain means being invariant under scaling operations in the frequency domain, and hence the phase spectrum is independent of the frequency (at least for a certain range of frequencies). This can be visualized by considering some addends of the Fourier series of a certain structure (see figure 6.8). In the example the signal is a step function at the position $x = 0$. The corresponding Fourier series consists of sine functions with odd frequencies. At the position $x = 0$ all sine functions are in phase, i.e., their phase is congruent. Hence, *phase congruency* can be considered as a measure for the existence of a structure *independently of the signal energy* [82, 83].

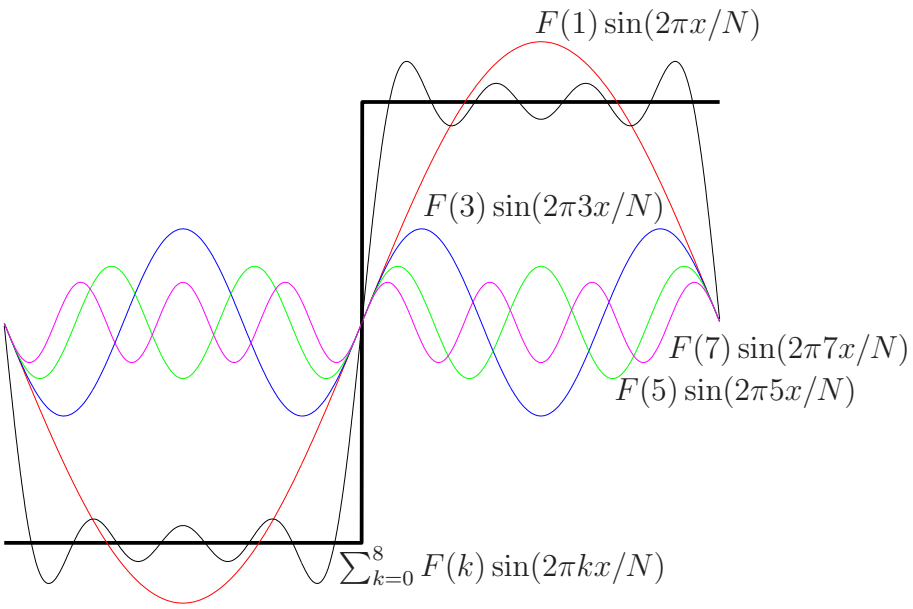

Fig. 6.8: Congruent phase at the position of the structure (here: step function).

The appropriate tool for estimating the phase for a local neighborhood is known to be a quadrature filter. Hence, those points where the local phase of the quadrature filter remains unchanged under scaling transformations are edge-points.

A sophisticated edge detection algorithm should discriminate clearly between points of congruency and others. A possible method is to calculate some filter responses for different scales and to define a certain function to estimate the congruency [82, 44, 121]. Another approach is to detect the zeros of the scale derivative of the quadrature filter. A zero in the scale derivative means that there is no change of phase for a small change of scale. The advantages of this method in contrast to those cited above are that the detection can be done at a certain scale (i.e., without taking into account neighbored scales) and that the positions of the zeros are easily detected with sub-pixel accuracy by linear interpolation.

6.2.2 Derivation of the Filter Set

The previous method can only be implemented if the s-derivative of the quadrature filter can be derived analytically. Fortunately, this is simple for the spherical quadrature filters consisting of the DOP filter $b(\mathbf{x})$ (4.38) and the DOCP filter $c(\mathbf{x})$ (4.43). The local phase is obtained according to (4.30) applied to the output of an

SQF, and therefore

$$\begin{aligned}\frac{\partial}{\partial s}\mathbf{e}_3 \arg((b+c)*\mathbf{f})^* &= \mathbf{e}_3^* \left\langle ((b+c)*\mathbf{f}\,\mathbf{e}_3)^{-1}\frac{\partial (b+c)*\mathbf{f}\,\mathbf{e}_3}{\partial s}\right\rangle_2 \\ &= \mathbf{e}_3^* \left\langle \frac{\mathbf{e}_3((b+c)*\mathbf{f})((b_s+c_s)*\mathbf{f}\,\mathbf{e}_3)}{|(b+c)*\mathbf{f}|^2}\right\rangle_2 \\ &= \mathbf{e}_3^* \frac{\langle(\mathbf{e}_3(b+c)*\mathbf{f})((b_s+c_s)*\mathbf{f}\,\mathbf{e}_3)\rangle_2}{|(b+c)*\mathbf{f}|^2} \\ &= \frac{(b*\mathbf{f})(c_s*\mathbf{f})+(c*\mathbf{f})(b_s*\mathbf{f})+(c*\mathbf{f})\wedge(c_s*\mathbf{f})}{|(b+c)*\mathbf{f}|^2}\mathbf{e}_3^* \ .\end{aligned} \tag{6.8}$$

The s-derivatives of the DOP filter b and the DOCP filter c are obtained in the frequency domain as

$$\begin{aligned}\frac{\partial B(\mathbf{u})}{\partial s} &= -2\pi|\mathbf{u}|\lambda^k \exp(-2\pi|\mathbf{u}|s\lambda^k) + 2\pi|\mathbf{u}|\lambda^{k-1}\exp(-2\pi|\mathbf{u}|s\lambda^{k-1}) \\ &= 2\pi|\mathbf{u}|\lambda^{k-1}(\exp(-2\pi|\mathbf{u}|s\lambda^{k-1}) - \lambda\exp(-2\pi|\mathbf{u}|s\lambda^k)) \qquad (6.9)\\ \frac{\partial C(\mathbf{u})}{\partial s} &= 2\pi\mathbf{u}I_2^{-1}\lambda^{k-1}(\exp(-2\pi|\mathbf{u}|s\lambda^{k-1}) - \lambda\exp(-2\pi|\mathbf{u}|s\lambda^k)) \ . \qquad (6.10)\end{aligned}$$

Somehow surprisingly, there are three distinct components in (6.8) that have to be checked: the first two terms in the numerator are bivectors containing $\mathbf{e}_3$ and the third term is an $\mathbf{e}_{12}$-bivector resulting from the wedge product. The latter component is easy to interpret as the generator of a rotation in the image plane: if the underlying structure is i1D, the orientation of the conjugate Poisson part $c*\mathbf{f}$ does not change, which means that its s-derivative $c_s*\mathbf{f}$ has exactly the same orientation as $c*\mathbf{f}$ itself. Hence the wedge product is only nonzero if the underlying structure is i2D, i.e., the orientation changes through scale. Unfortunately, this outer product cannot be used for stable estimation of the intrinsic dimension since it is highly affected by noise. On the other hand, it can be used as a confidence measure of the edge detection since any large value in that component means that the vector part is not very accurate, either caused by noise or by an i2D structure.

The first two bivectors can be interpreted as the generators of a pure phase change, and they must therefore be zero at a line or an edge. For estimating the exact position of the line / edge, the zeros of that bivector part are estimated with subpixel accuracy. This is done by projecting the two components onto the local orientation and subsequent fitting of a 1D linear function of limited extent (see e.g. [15, 39]). In order to have an efficient implementation, this optimization is done by a normalized convolution. Only those zeros where the slope of the fitted linear function is larger than a fixed certain threshold (0.5 in all cases) are kept for the edge image. Note that no hysteresis thresholding is applied.

6.2.3 Experiments and Comparisons

The results of the presented method are visualized in figure 6.9. All images are taken from the website [105]. In order to have an idea of the performance, the new approach is compared to edge detection algorithms taken from the image processing toolbox of Matlab: the Sobel edge detector (see also figure 6.9), the Laplacian of Gaussian (LOG) edge detector, and the Canny edge detector [28] (see both in figure 6.10). Note that the algorithms of the Matlab toolbox do much more than just convolving the corresponding filter masks and applying simple thresholds. Furthermore, I compare my approach with the phase congruency edge detector of Kovesi [82]. The parameters of all algorithms can be found in table 6.2.

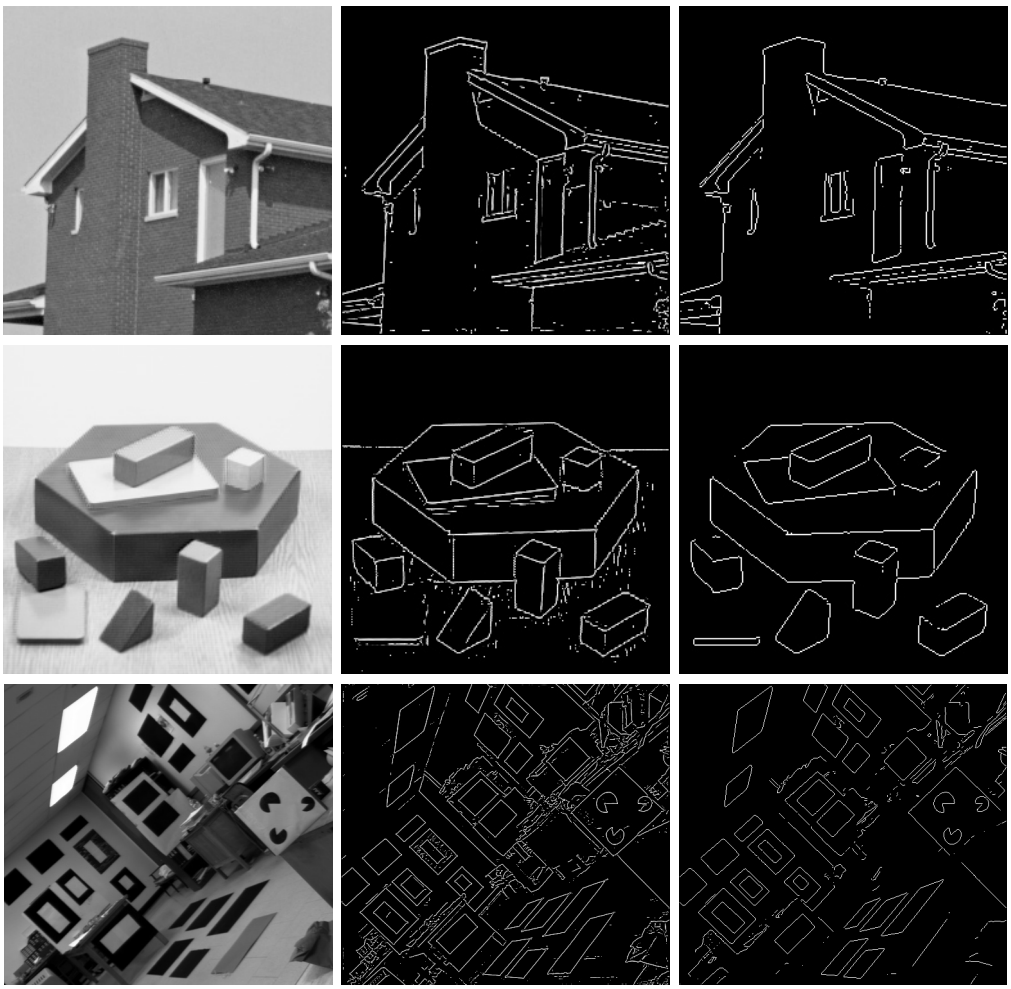

Fig. 6.9: Edge detection of test images. From top to bottom: 'house', 'blox', and 'lab'. From left to right: original image, edge detection using the presented method, and output of the Sobel edge detector.

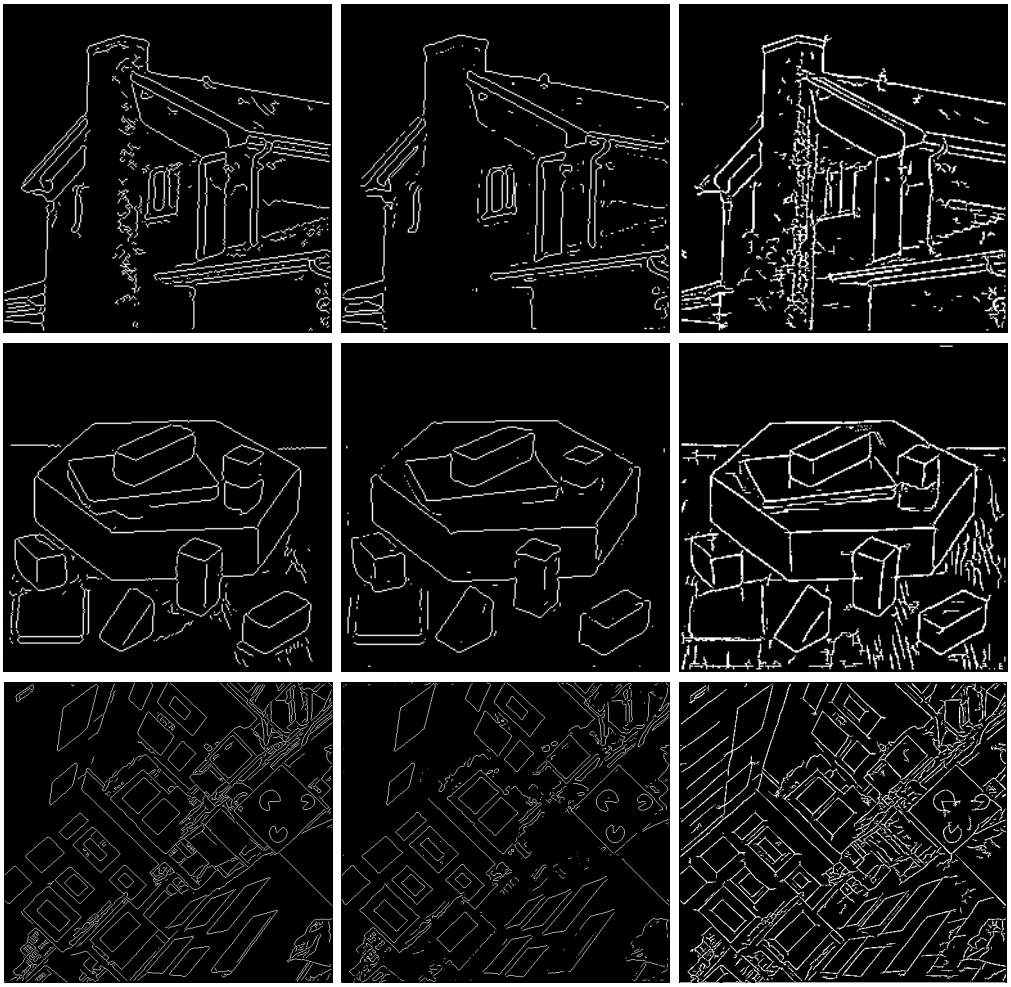

Fig. 6.10: Edge detection of test images. From top to bottom: 'house', 'blox', and 'lab'. From left to right: edge detection using the Canny detector, the Laplacian of Gaussian edge detector, and phase congruency (Kovesi).

All parameters are standard values or automatically generated by the Matlab algorithms in order to obtain a comparison of fully automatic approaches. The new method and the method of Kovesi are applied with constant parameters. The new method is the only one which directly yields subpixel accuracy, which is indicated by a grayscale coding. Each one of the applied methods has its own (subjective) qualities and I summarize these properties in table 6.3.

The property 'noise' stands for the noise present in the edge image. Very good means that no noise is contained in the output. False positives are detected edges where no edge is present in the image. Sometimes it is hard to distinguish these from noise in the output. Very good means that no artificial edge is created. False

approach	'house'	'blox'	'lab'
new	$s_1 = 1$ $s_2 = 2$ $t = 0.5$	$s_1 = 1$ $s_2 = 2$ $t = 0.5$	$s_1 = 1$ $s_2 = 2$ $t = 0.5$
Sobel	$t = 26.3$	$t = 22.4$	$t = 22.7$
Canny	$t_1 = 0.0312$ $t_2 = 0.0781$ $\sigma = 1$	$t_1 = 0.0250$ $t_2 = 0.0625$ $\sigma = 1$	$t_1 = 0.0312$ $t_2 = 0.0781$ $\sigma = 1$
LOG	$t = 1.43$ $\sigma = 2$	$t = 1.48$ $\sigma = 2$	$t = 1.26$ $\sigma = 2$
Kovesi	$r = 1.5$ $t_1 = 0.15$ $t_2 = 0.30$	$r = 1.5$ $t_1 = 0.15$ $t_2 = 0.30$	$r = 1.5$ $t_1 = 0.15$ $t_2 = 0.30$

Tab. 6.2: Parameters of the different approaches for edge detection. All parameters are either standard values (new approach, Kovesi, the σ for Canny and LOG) or chosen automatically (the remaining parameters).

negatives are edges which are not detected. Very good means that no edge is missing. The behavior for i2D points is bad if errors occur at junctions, e.g. corners are rounded or artificial structures are created. Very good means that corners stay sharp and no additional structures are created.[2] The connectivity indicates if detected edges are connected or not. Very good means that no edge is unconnected. However, there are some characteristic properties which cannot be expressed by the table.

- The new approach behaves quite flexible, though it could be less noise sensitive. It misses some edges and destroys the connectivity in some places. All three effects could be reduced if the slope threshold would be replaced by a hysteresis threshold (as it is used for the approach of Kovesi). Note that this method also detects lines.
- The Sobel operator tends to detect too few edges. Hence it creates no false positives, but has a lot of false negatives. This may be caused by the automatic parameter estimation which might be too restrictive.
- The Canny edge detector tends to detect textures as edges (see e.g. house image). Furthermore, some corners become very round.
- The Laplacian of Gaussian behaves similar to the Sobel operator concerning false positives and false negatives. Additionally, it tends to smooth corners.

[2] One cannot expect an edge detector to do more than detecting edges. However, this should not be disturbed by i2D points.

method	property	'house'	'blox'	'lab'
new	noise	+	+	+
	false-positives	++	++	+
	false-negatives	+	++	+
	i2D behavior	+	+	+
	connectivity	+	+	++
Sobel	noise	++	++	++
	false-positives	++	++	++
	false-negatives	-	-	-
	i2D behavior	+	+	+
	connectivity	+	+	+
Canny	noise	○	+	++
	false-positives	-	○	++
	false-negatives	+	+	+
	i2D behavior	○	+	+
	connectivity	++	++	++
LOG	noise	+	++	+
	false-positives	+	++	++
	false-negatives	-	-	-
	i2D behavior	-	○	○
	connectivity	+	○	+
Kovesi	noise	-	○	○
	false-positives	--	-	-
	false-negatives	+	+	++
	i2D behavior	--	-	--
	connectivity	+	+	++

Tab. 6.3: Subjective quality of the applied edge detectors for the different test images. The quality is assessed by five grades: ++ means very good, + means good, ○ means average, - means poor, and -- means bad.

- The algorithm of Kovesi behaves opposite to the Sobel operator and the LOG operator. It is noise sensitive and detects too many edges. Furthermore, it tends to create artifacts at corners. However, it also detects lines.

Surely, there are other gradient based detectors (e.g. Roberts edge detector) and other sophisticated methods for gradient estimation (e.g. [99]), but I do not want to stress this comparison too much since edge detection is only one possible application of the whole theory of the monogenic signal. Another possible edge detector using this approach is obtained by detecting local maxima in the local amplitude of the SQF responses, but this method does not behave as stable as the phase-based approach.

6.3 Stereo Correspondence and Disparity Estimation

The responses of SQFs can also be used for the matching problem which occurs in the context of stereo correspondence and for estimating the disparity between two similar images. Both approaches are briefly sketched in the following.

6.3.1 Stereo Correspondence by Matching on the Epipolar Line

The stereo algorithm has been developed in cooperation with N. Krüger and M. Pörksen [117]. It is based on an older algorithm proposed by N. Krüger and M. Ackermann [89, 90]. The idea is the following. By means of phase congruency edges are detected in both images. The edges in the left image are subsampled in order to reduce the data. For every remaining edge point the epipolar line in the right image is constructed. For all edge points in the right image which lie on the epipolar line, a non-Euclidean distance to the left edge point is calculated. The point on the epipolar line with the least distance is used as the corresponding point if the distance is below a certain threshold. In contrast to the local processing in this approach, the older method used the intersection of Hough lines and the epipolar line to detect the corresponding point.

The distance measure plays the central role in the new algorithm. It is based on SQF responses and tries to unify different local signal properties in one scalar measure, representing the similarity of the local neighborhoods. It is based on four components: the phase congruency in both edge points, the phase difference between the two images, and the orientation difference between the two images. From several experiments it is known that the phase of the same physical point in both images is nearly the same (unless occlusion occurs). Furthermore, experiments have been made concerning the difference of the 2D orientations in both images depending on the 3D spatial position of a line. From these results a distance measure d on the space of phase difference and orientation difference has been constructed, which is adapted to the variances of the two estimates:

$$d_1 = \sqrt{\alpha \,\mathrm{angdist}(\varphi_{\mathrm{left}}, \varphi_{\mathrm{right}})^2 + (1-\alpha)\mathrm{angdist}(\theta_{\mathrm{left}}, \theta_{\mathrm{right}})^2} \tag{6.11}$$

$$d_2 = \sqrt{\alpha \,\mathrm{angdist}(\varphi_{\mathrm{left}}, -\varphi_{\mathrm{right}})^2 + (1-\alpha)\mathrm{angdist}(\theta_{\mathrm{left}}, \theta_{\mathrm{right}} + \pi)^2} \tag{6.12}$$

$$d = \min(d_1, d_2) \ , \tag{6.13}$$

where $\mathrm{angdist}(\gamma, \beta)$ indicates the smaller angle between γ and β, and α is a weighting constant which is has been estimated in the experiments (see above). The distance measure d is to some extent similar to a Mahalanobis distance, see also section 6.1 and [9]. By means of this distance, the local grayscale information of the structures is compared. If the grayscale structure in the two points is similar (e.g. an edge

with positive slope in a certain direction), the distance is very small. Taking some reciprocal function of the distance yields a similarity measure. This is additionally multiplied with the two phase congruencies, pc_{left} and pc_{right}, which indicate the probability of the presence of a structure. The result is a probability measure that two points are the projections of the same physical point:

$$p = pc_{\text{left}} pc_{\text{right}} (1 - d/\pi) \ . \tag{6.14}$$

The advantage of this method is that it combines some ideas from photometric matching with properties from feature based matching. The local phase describes to some extend the grayscale structure. It is invariant with respect to shifts and rotations in the image plane. Furthermore, the local phase remains constant if a dilation is performed, unless the passband of the SQF is left. For i1D structures, this is also true for 2D affine transforms up to a certain degree.

The local orientation however, directly reflects 2D rotations and is affected by 2D shears. Using the knowledge of the projection matrices, it is possible to calculate the 3D orientation from the two 2D orientations. On the other hand, the orientation difference between the two images is small in most cases (at least for large distances and non-singular 3D orientations) and can therefore be used for the matching. Finally, the edge detection reduces the computational load of the algorithm, because it reduces the data to be processed.

Using a combination of the features phase, orientation, and phase congruency in the matching is to some extent comparable to the invariant matching resulting e.g. from canonical coordinates [122, 130, 131]. Obviously, the features resulting from an SQF are not *complete* in the mathematical sense, but in the case of stereo correspondence they are *sufficiently discriminative*. In figure 6.11 one example of 3D reconstruction based on the presented stereo correspondence algorithm is shown.

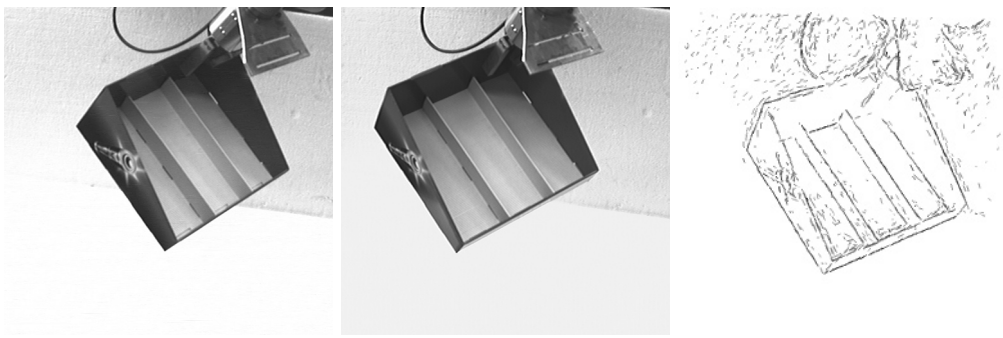

Fig. 6.11: Example for 3D reconstruction from stereo. From left to right: image from left camera, image from right camera, and representation of corresponding line segments which have been found.

6.3.2 Disparity from Monogenic Phase

A second application which is closely related to stereo vision is the problem of disparity estimation. In [67] Hansen uses phase information to recover local shifts in order to estimate the disparity in stereo images. This is possible since a shift in the spatial domain yields a linear phase factor for the analytic signal. Unfortunately, the classical quadrature filters yield systematic phase distortions if the shift is directed differently than the quadrature filter. This is no longer true for the SQF responses which yield a linear phase in *every direction* and also represent the line of projection in the case of i1D structure. In the latter case, namely, a shift parallel to the structure cannot be recovered from phase information, which is a problem in principle (the aperture problem) and not specific to phase-based approaches. Hence, the shift that is obtained from the phase is the true shift projected onto the normal of the structure. However, knowing the normal shift and the projection orientation is necessary to recover the true shift (e.g. by segmentation and interpolation).

A well suited method for estimating the local shifts is to apply a multi-scale approach. On a coarse scale, large movements of large objects can be detected. Using this information on a finer scale, the shift on the coarse scale can be compensated and a more accurate estimate of the shift is obtained from the phase information. Taking into account finer and finer scales, a disparity map with subpixel accuracy can be obtained.

If the product of the center frequency of the involved SQF and the shift in the spatial domain is small enough, the modulation which results from the shift represents the latter uniquely. For the purpose of disparity estimation, the SQF responses for the left and the right image are evaluated, yielding $\mathbf{f}_\mathrm{l}$ and $\mathbf{f}_\mathrm{r}$. The relative modulation term of the two responses represents the relative shift. The modulation term s is obtained from $\mathbf{f}_\mathrm{l}$ and $\mathbf{f}_\mathrm{r}$ according to

$$s = \mathbf{f}_\mathrm{l}\mathbf{f}_\mathrm{r}^{-1} = (\mathbf{f}_\mathrm{l} \cdot \mathbf{f}_\mathrm{r} + \mathbf{f}_\mathrm{l} \wedge \mathbf{f}_\mathrm{r})|\mathbf{f}_\mathrm{r}|^{-2} \quad . \tag{6.15}$$

The modulation term s is a spinor, and hence applying the logarithm of $\mathbb{R}_3^+$ yields a bivector which represents the difference of the monogenic phase of the two SQF responses. This phase difference bivector is basically proportional to the relative shift between the two image regions. The ratio between the phase vector and the shift is given by the center frequency. Note that the shift which is estimated in this way is the *normal shift* and not the true shift (due to aperture).

In the special case of small phase differences, i.e., less than 0.2 radians, the logarithm can be identified with the bivector part of its argument. This corresponds to a first order Taylor approximation. The shift $\Delta\mathbf{x}$ is then proportional to the wedge product of the two SQF responses: $\Delta\mathbf{x} \sim \mathbf{f}_\mathrm{l} \wedge \mathbf{f}_\mathrm{r}$. The estimation of the normal shift is especially fast in this case; two filters have to be applied and one multiplication

has to be performed. This method can also be used to estimate the normal flow in an image sequence, it is not restricted to disparity estimation.

In a first experiment, a synthetic i1D pattern is moving on a circle. The normal flow of this sequence is estimated by the described algorithm and the results are illustrated in figure 6.12. The modulus of the estimated normal flow can be found in figure 6.13, and the orientation of the flow is constant since the orientation of the texture is not changed. Obviously, the normal flow is estimated correctly, because the projection onto the normal vector of the texture yields a cosine of the orientation difference between texture orientation and motion direction.

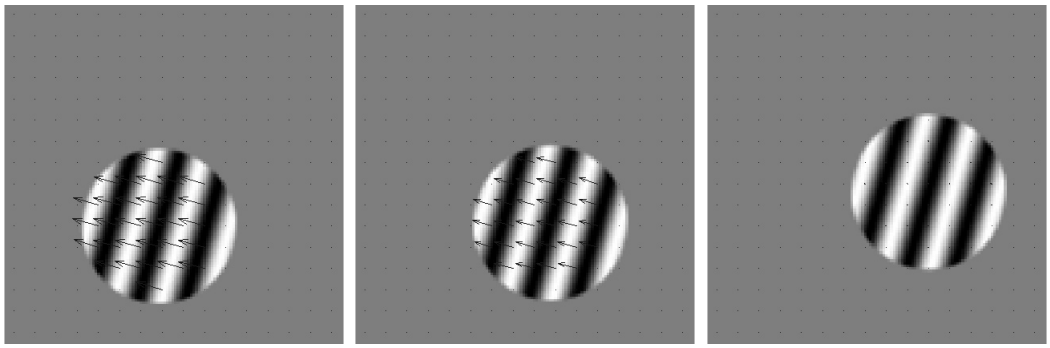

Fig. 6.12: Flow field estimation of a synthetic sequence that shows a disc which moves on a circle (origin in the center of the image). The presented method estimates the normal flow, which varies from the true flow (left image) to zero flow (right image). The flow is weighted by the cosine of the angle between motion vector and the local orientation, see image in the center.

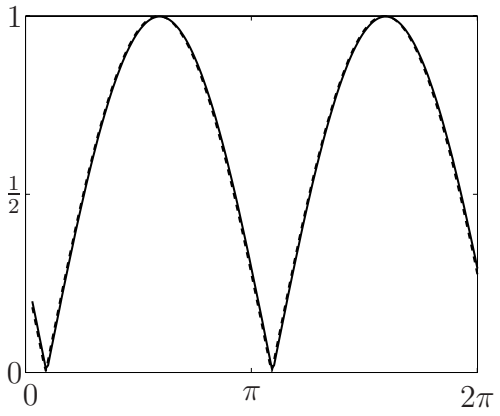

Fig. 6.13: The modulus of the normal flow for an i1D pattern. The ideal curve is indicated by the dashed line and the estimated curve is indicated by the solid line.

For i2D patterns, it should be possible to estimate the true flow. In figure 6.14 the same moving disc from figure 6.12 can be found, but with an i2D texture. For this sequence it is possible to estimate the correct flow direction (see figure 6.15, right). However, the modulus of the flow is not estimated correctly, since the monogenic signal shows amplitude distortions for i2D signals. Nevertheless, the flow modulus remains constant (see figure 6.15, left).

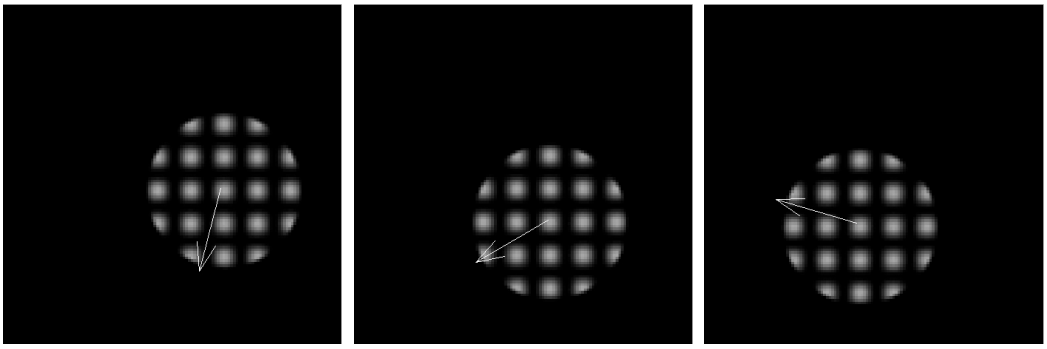

Fig. 6.14: Flow field estimation of a synthetic sequence with an i2D texture. The presented method allows to estimate the correct direction of the flow, although the modulus is wrong (see text).

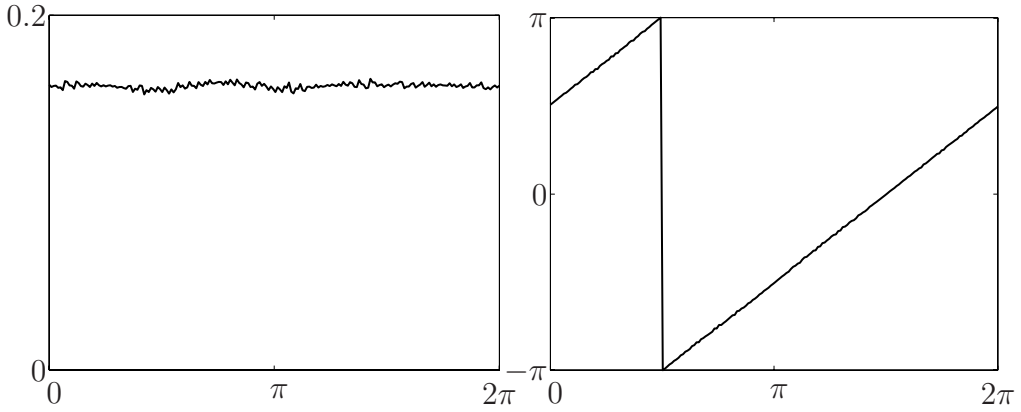

Fig. 6.15: Estimated flow. Left: the modulus of the flow is nearly constant, but it is depending on the texture. Right: the direction of the flow is estimated correctly.

The presented algorithm for the motion field and disparity estimation is only a preliminary version which should show that it works in principle. For the correct i2D motion and disparity estimation it is necessary to extend this idea to the structure multivector. Furthermore, efficient multi-scale implementations are not available yet.

6.4 Adaptive Smoothing

In section 4.5 I have mentioned that spherical quadrature filters can be used to detect local maxima in scale-space, which can be interpreted as the *intrinsic scales* of an image. For the application of denoising images, the following algorithm has been developed (see also [48]).

6.4.1 The Adaptive Smoothing Algorithm

1. For every position detect the first local maximum of the local amplitude descending through scale-space from fine to coarse. These maxima form the scale index.

2. Due to theoretic considerations the scale index must be a smooth function since the scale cannot change arbitrarily between adjacent pixels. Therefore, outliers are eliminated using a median filter.

3. For every position choose the Poisson *lowpass* filter output which corresponds to the smoothed scale index.

The actual algorithm uses a coarsest level in scale-space for reducing the search-space. This coarsest level s_{M} is chosen such that a further smoothing would average too large parts of the image (e.g. a quarter of the image). The scale index is obtained by the function

$$\mathrm{index}(\mathbf{x}) = \begin{cases} s_{\mathrm{m}} & \text{if } \exists\, \varepsilon > 0 : |\mathrm{SQF}(\mathbf{x}, s_{\mathrm{m}})| \geq |\mathrm{SQF}(\mathbf{x}, s)| \quad \forall s < s_{\mathrm{m}} + \varepsilon \\ & \quad \wedge \quad s_{\mathrm{m}} < s_{\mathrm{M}} \quad \wedge \quad |\mathrm{SQF}(\mathbf{x}, s_{\mathrm{m}})|^2 \geq E_{\mathrm{T}} \\ s_{\mathrm{M}} & \text{else} \end{cases} \tag{6.16}$$

where the coarsest scale s_{M} is chosen if the maximal energy is lower than a threshold E_{T}. The energy at a certain scale s is estimated by the (square of) the local amplitude of the SQF, $\mathrm{SQF}(\mathbf{x}, s) = b_{s_{\mathrm{M}},2,\mathrm{ld}(s)}(\mathbf{x}) + c_{s_{\mathrm{M}},2,\mathrm{ld}(s)}(\mathbf{x})$.

The algorithm is motivated by the fact that the response of the SQF shows local maxima when it is applied in the 'correct' scale, where in general more than one maximum exist, see figure 4.17 on page 99. The threshold of the energy at the maximum suppresses weak textures by setting the index function to maximal smoothing.

In the second step, median filters are applied to the index function since there might occur outliers in the maximum detection. Due to the spatial extent of the SQF responses, the scale index must be smooth to some degree, and therefore it is not

reasonable to have an arbitrarily changing scale index. An example for a median-smoothed scale index can be found in figure 6.16. It becomes quite clear from this example that the scale index is fine at areas where structure is present and it is coarse where no structure is present.

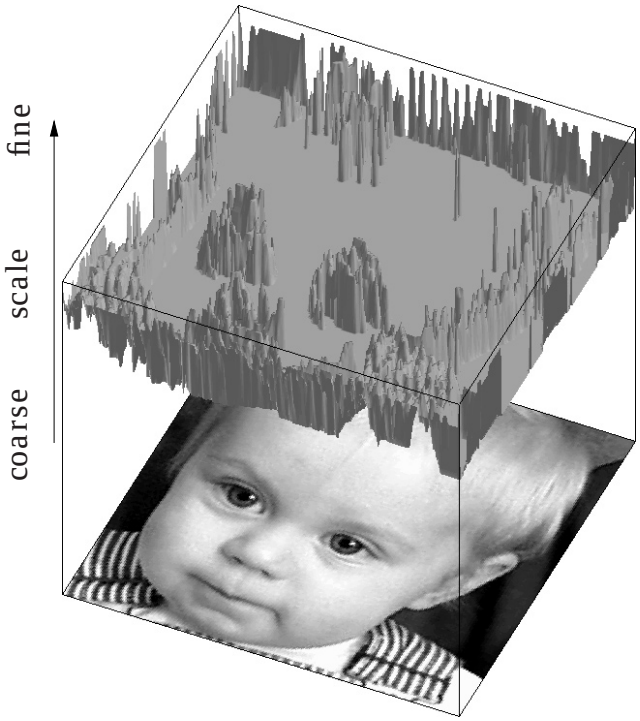

Fig. 6.16: Image and corresponding scale index surface in scale-space.

The scale index is then used to obtain a denoised image by taking at each position the Poisson lowpass filter output at the appropriate scale. The whole method is therefore a nonlinear, adaptive, and isotropic smoothing. Since the method is not derived from a partial differential equation as it is the case for most diffusion approaches, it does not necessarily preserve the (local) DC-level, see also [143].

6.4.2 Experiments and Comparison

In the first experiment, the algorithm has been applied to two test images (see figure 6.17). The images have both been processed with the same threshold and the same maximal scale. In both cases a 3×3 median filter has been applied to the scale index obtained from (6.16). The dark shadows at the top and left border of the smoothed house image are artifacts resulting from performing the convolutions in the frequency domain. Obviously, textures with low energy are removed while

edges and other structures with high energy are preserved. Whether a structure is preserved or not depends on the choice of the threshold E_{T} which sets the minimum energy of a structure to be kept. In the case of noisy images, this energy threshold can be estimated according to a Rayleigh distribution (see e.g. [83]).

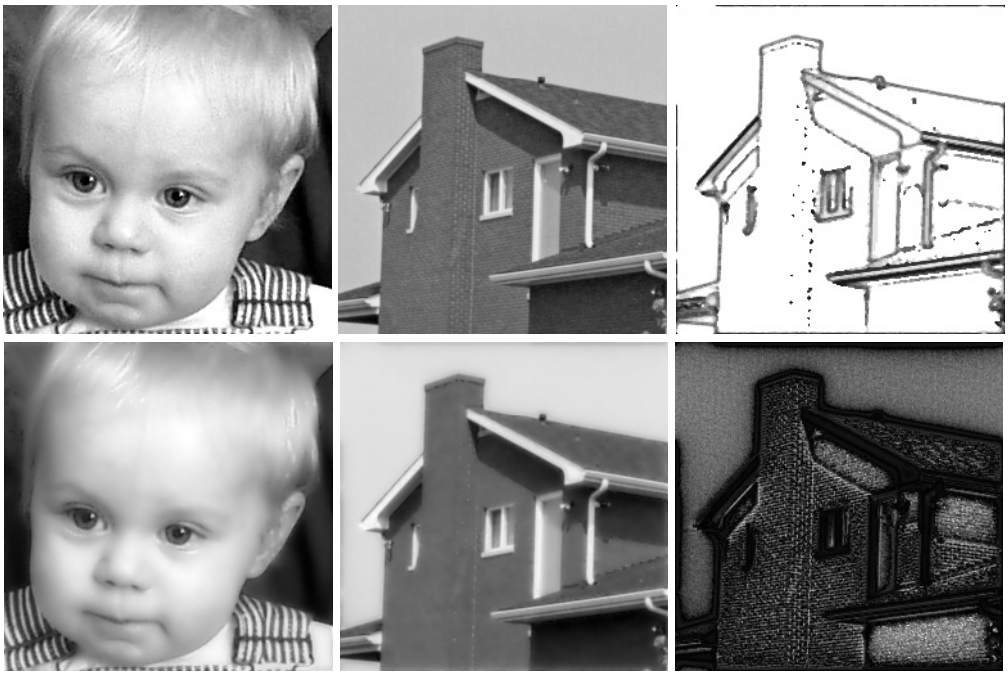

Fig. 6.17: Upper row from left to right: original image 'child', original image 'house', scale index of 'house'. Bottom row from left to right: adaptively smoothed image 'child', adaptively smoothed image 'house', relative absolute difference between original and smoothed version of 'house'.

In the second experiment, Gaussian distributed white noise is added to the first image in figure 6.17 (see figure 6.18, left) and a standard method of isotropic diffusion has been applied to it. The used method[3] has been proposed in [29]. The result can be found in the center image of figure 6.18. The typical behavior of isotropic diffusion can be seen well: near the edges the noise is still present. The new approach results in the image on the right of the same figure. Except for some singular points the noise has been removed completely. The energy threshold has been chosen by the mean energy of the SQF output on the finest scale (see [83]). In contrast to the diffusion algorithm, the presented method tends to over-smooth slightly those region boundaries where the energy of the edge is lower than the energy of the noise (see e.g. upper left corner of the right image of figure 6.18).

[3] Thanks to J. Weickert for providing the code.

Fig. 6.18: Test image with noise (left), variance: $6 \cdot 10^3$. Isotropic diffusion (parameters: contrast $\lambda = 3$, noise scale $\sigma = 2$, time step $\Delta t = 0.2$, and 50 iterations) yields the result in the middle, variance: $4.5 \cdot 10^3$. The image on the right shows the result of the presented approach, variance: $3.5 \cdot 10^3$.

The algorithm can also be used to sparsify images. Instead of constructing the denoised image from lowpass filtered versions using the scale index, the filter response of the monogenic signal is kept for local maxima of the energy only. That means that for the 'else' case in (6.16) the result is set to zero (nothing is stored). For the example of the house image, the amount of the remaining information can be estimated by the dark areas in the index image in figure 6.17, which cover about 20% of the image plane. The representation which is obtained includes information of the monogenic signal only at positions with high local energy at an appropriate scale.

The scale-adaptive smoothing can be further improved by using the method of local metrics (see section 5.1). The resulting approach then performs an anisotropic, nonlinear smoothing. Furthermore, the fixed threshold E_{T} can be replaced by some more sophisticated method.

6.5 Corner Detection and Curvature Estimation

According to the headline, the topic of this section is corner detection (see e.g. [106, 68]) and curvature estimation (see e.g. [4]). The key to these fundamental problems of 2D signal processing is the minor amplitude of the structure multivector. Hence, this is the first application of the i2D extension of the monogenic signal.

6.5.1 The Local Isotropy and Curvature

The local isotropy can be estimated according to (5.71) which also gives an idea of the local intrinsic dimension. The local isotropy is a value between zero and one, where zero indicates an i1D structure and one indicates an i2D structure (assumed that the total amplitude is non-zero). Corner detection is basically a detection of centers of locally i2D neighborhoods. However, the local isotropy is not directly suited for corner detection since (uncorrelated) noise is always i2D, and therefore all regions in the image which are dominated by noise would be detected as corners. Thus, it is reasonable to introduce some energy measure. A possible choice is $A_{\mathrm{m}}(\mathbf{x})$, the energy of the minor structure, which can also be considered as the product of the local isotropy with the square root of the two local amplitudes, see (5.71).

To show that the minor amplitude is actually an appropriate mean for detecting i2D points, I have applied the SMV to an image showing several line stoppings with different orientations, see figure 6.19. Obviously, the local amplitudes are independent of the orientation. Furthermore, the minor amplitude is about half of the major amplitude, which I have tried to visualize by using the same range for both amplitude images.

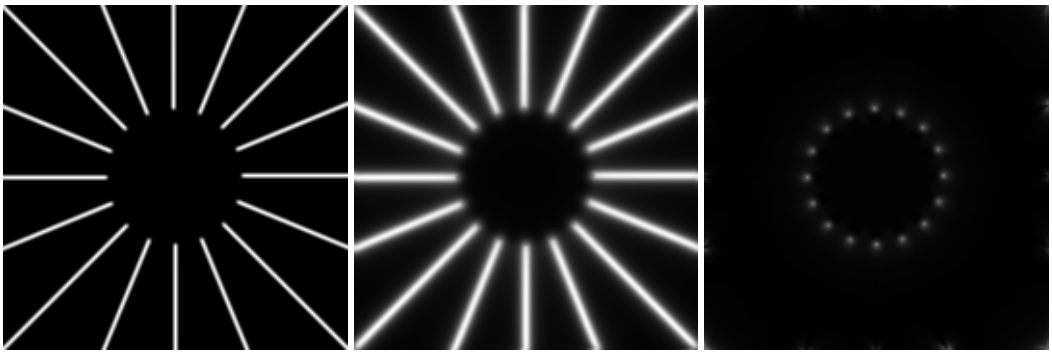

Fig. 6.19: From left to right: original image showing line stoppings, the major amplitude, and the minor amplitude (in the same range as the major amplitude).

Curvature is somehow related to the intrinsic dimension since it can be seen as a continuous measure of the 'cornerness' of an edge or line. Hence, it is reasonable to assume a relation between curvature and the local isotropy. The following two experiments show that there is actually a relation. In the first experiment (see figure 6.20), the local isotropy is calculated for a logarithmic spiral[4] modulation. Obviously, the local isotropy is *constant*, except for the center and the image borders. This can also be verified by integrating over the angular coordinate, which gives the

[4] A logarithmic spiral is defined by having a curvature which is reciprocal to the radius.

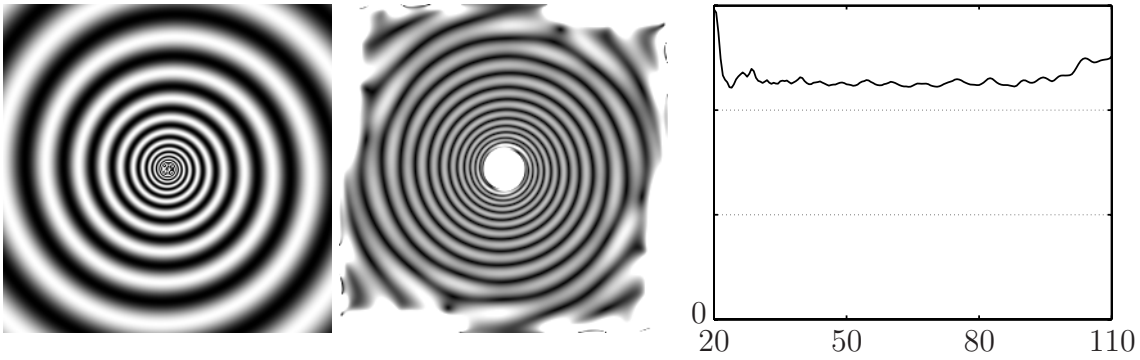

Fig. 6.20: From left to right: logarithmic spiral modulation, local isotropy, and radial plot of isotropy.

radial plot on the right of figure 6.20. The consequence of this experiment is the following: the local isotropy is *invariant* under similarity transformations, since the logarithmic spiral modulation consists solely of similar modulation patterns. As a consequence, the local isotropy decreases if the modulation frequency increases and the curvature remains the same and vice versa. In other words, the local isotropy can be considered as a scale-relative measure of curvature, i.e., the difference of the curvatures of two structures, which are on the same scale, is represented by the isotropy. This is verified by a simple experiment, see figure 6.21. The test image

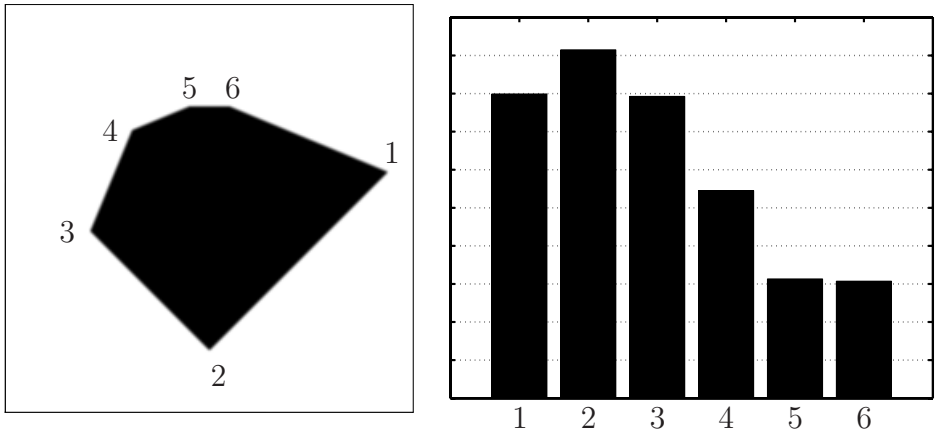

Fig. 6.21: Left: image showing corners of various angles. Right: minor amplitudes at the corners.

shows six corners which are on about the same scale. The minor amplitudes at these corners yield estimates for the deviation from a rectangular corner. Whether the corner angle is less or more than $\pi/2$ can be obtained from the two phases at the corner. In order to develop a sophisticated method for estimating the curvature in a quantitative way, further investigations are necessary. However, it is possible to use the curvature measure for a corner detection algorithm.

6.5.2 Corner Detection Experiments and Comparisons

The implemented algorithm for corner detection searches for local maxima of $A_{\mathrm{m}}(\mathbf{x})$. A local maximum is indicated whenever a signal value is larger than all other signal values with a distance of less than eight pixels. If the energy of the local maximum is larger than half the mean amplitude of the total amplitude, the local maximum is supposed to be a corner. The two parameters (the radius of the search for a local maximum and the threshold) are chosen without any special preference and may be suboptimal for most images. More sophisticated methods of parameter estimation surely yield better results. However, all the subsequent experiments have been performed with these parameters. The radial bandpass which is combined with the structure multivector is the same as in section 6.2.

In figure 6.22 the output of the presented method for three different images can be found. In order to compare them with some standard approaches, the results of the Kitchen-Rosenfeld (KR-) detector, the Plessey (Harris-Stephens, HS-) detector, the curvature-scale-space (CSS) approach, and the SUSAN-detector are also presented in figure 6.22 and in figure 6.23. The images for the latter four approaches are taken from [105].

Actually, it is not fair to compare single-scale approaches like the presented one or the Plessey detector with a multi-scale approach like the CSS-approach, but since the images are taken from a website about the CSS-approach, I included them although. However, one should keep in mind that some of the other detectors can also be applied in multiple scales. Moreover, the CSS-approach uses the most sophisticated postprocessing on the basis-filter responses, combining edge-information in the semi-local neighborhood [106]. Accordingly, the CSS-detector shows the best performance.

In the subsequent table 6.4, I compare the results of the different approaches according to the criteria false-positives, false-negatives, and localization accuracy. Obviously, each of the three images has its own difficulties. The localization accuracy is unproblematic; it is without exception good or at least sufficient. This is different with the false classifications. In the 'lab' image, the HS-detector shows a lot of false-positives along lines and fails to detect a lot of the quite 'clear' corners of the rectangles. These 'clear' corners also provide problems to the SUSAN-detector and partially to the KR-detector which also shows some false-negatives.

For the 'blox' image, the results are basically better with the exception of the SUSAN-detector which shows a significant number of false positives. For the 'house' image, the HS-detector shows again a lot of false positives along lines and fails to detect some corners. The presented method and the KR-detector also fail to detect all corners. For a more detailed comparison the reader should refer to the images. Basically, the results of the corner detection are surprisingly good compared to the edge detection; they are of a similar quality.

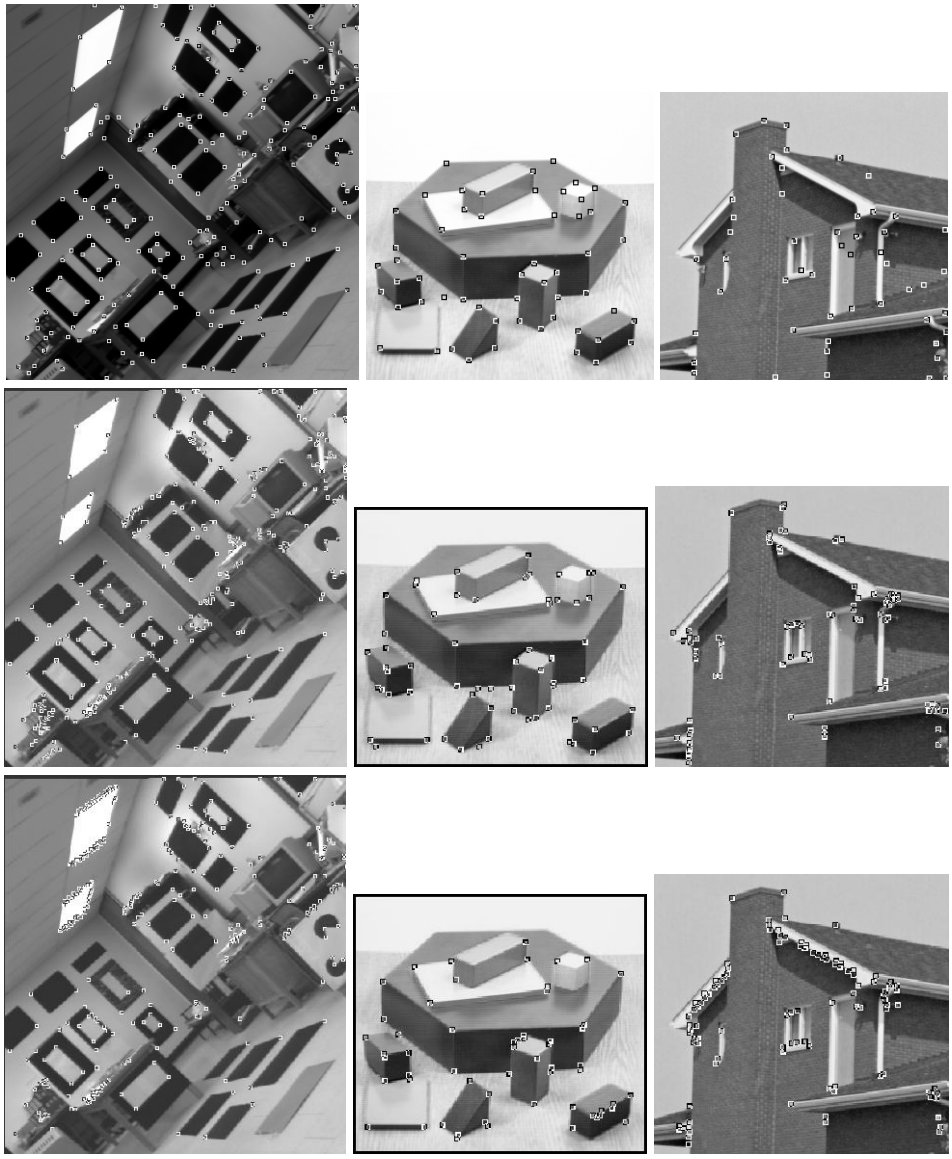

Fig. 6.22: From left to right: detected corners in the 'lab' image, the 'blox' image, and the 'house' image. From top to bottom: output of the presented approach, of the Kitchen-Rosenfeld detector, and of the Plessey (Harris-Stephens) detector.

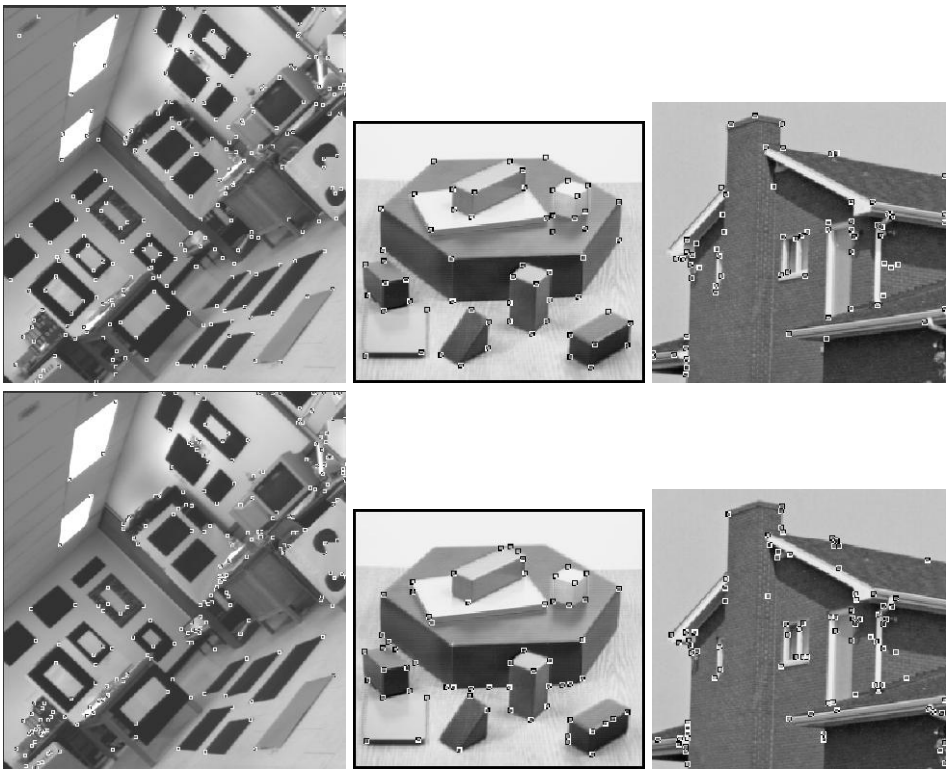

Fig. 6.23: From left to right: detected corners in the 'lab' image, the 'blox' image, and the 'house' image. Upper row: output of the curvature-scale-space approach. Bottom row: output of the SUSAN-detector.

Whereas the results of edge detection by monogenic phase congruency are already satisfactory, the corner detection based on the structure multivector still shows a need to be improved. Besides the detection of local maxima which should be replaced by more sophisticated methods as polynomial fitting, local gain control etc., the filtering itself can be improved. Applying the methods of local metrics and of adaptive scale selection, the filter responses probably become clearer, i.e., the local maxima become sharper. Implementations of these methods will follow in some future work and are out of the scope of this thesis.

6.6 Summary of Chapter 6

In this chapter several applications of the proposed theoretic framework have been presented: orientation estimation, edge and corner detection, adaptive smoothing, stereo correspondence, and disparity estimation. The distinct applications differ with respect to their current status, i.e., some can already be applied in real systems whereas others are still to be further developed.

method	property	'lab'	'blox'	'house'
new	false-positives	+	+	+
	false-negatives	+	++	○
	localization	+	+	○
KR	false-positives	○	+	○
	false-negatives	○	○	○
	localization	+	+	○
HS	false-positives	-	○	-
	false-negatives	-	○	○
	localization	+	+	○
CSS	false-positives	++	++	+
	false-negatives	+	++	+
	localization	+	+	+
SUSAN	false-positives	+	○	○
	false-negatives	-	○	+
	localization	++	+	+

Tab. 6.4: Subjective quality of the applied corner detectors for the different test images. New means the presented method, KR means Kitchen-Rosenfeld detector, HS means Harris-Stephens (Plessey) detector, CSS means curvature-scale-space approach, and SUSAN means the SUSAN-detector. The quality is assessed by five grades: ++ means very good, + means good, ○ means average, - means poor, and -- means bad.

The orientation estimation section includes considerations of noisy data, yielding an optimal orientation estimation which has already been proposed in chapter 5. The edge and line detection by the implicit scale-derivative of the phase is pretty efficient and powerful, although it might be further optimized with respect to automatic parameter estimation. The corner detection which is based on the minor amplitude works sufficiently good for a single scale approach. It should be extended to a multi-scale or semi-local associative approach. The remaining applications still need further improvement. The adaptive smoothing is based on searching for local maxima in the isotropic scale. This should be extended to anisotropic scales and a PDE-based algorithm. The stereo correspondence algorithm is based on searching the epipolar line for a similar SQF response. This could also be improved by taking into account several scales and the full structure multivector. Finally, the disparity from monogenic phase is still a sketch for an application, although the idea to estimate the displacement vector by means of the local isotropic phase is quite attractive.

Chapter 7

CONCLUSION

> *'The last thing one knows when writing a book is what to put first.'*
>
> Blaise Pascal (1623-1662)

This chapter is quite short, because I have given a short summary at the end of each chapter. However, the subsequent section provides an overview of the results. The final section of this chapter deals with multidimensional extensions of the presented methods, open problems, and future work.

7.1 Summary of the Results

The main outcome of this thesis can be summarized as follows:
It is possible to extend the 1D analytic signal and 1D quadrature filters to 2D in an isotropic way. This extension is not just a 2D analytic signal but yields a whole new framework of 2D low-level signal processing.

The 2D generalization of the 1D analytic signal consists of a sophisticated multivector-valued signal, the structure multivector. It includes orientation dependent factors which can be used to estimate the local orientation. Furthermore, it includes major and minor 2D amplitudes which directly correspond to the two eigenvalues of the structure tensor and can be used to estimate the local intrinsic dimension. Last but not least it contains two phases in the classical sense which describe the structures of two superposed perpendicular i1D signals. In the case of i1D signals, the structure multivector reduces to a 3D vector-valued signal, the monogenic signal. The latter contains solely the amplitude, orientation, and phase information of one i1D signal and it is obtained from the structure multivector by setting the isotropy to zero.

Besides the generalization of quadrature in 2D, the mathematical framework also yields a new linear scale-space which can be used to define radial bandpass

filters. These filters constitute a continuous wavelet transform and allow to define quadrature filters based on the monogenic signal and the structure multivector. Introducing local metrics, the radial symmetric filters are changed to elliptic filters which improves the stability of the orientation and phase estimation under the presence of noise. Furthermore, the local metric allows to create scale and isotropy adaptive filters which behave appropriately to the intrinsic dimension and main orientation of the local structure of the signal.

The generalized quadrature filters based on the structure multivector and the mentioned continuous wavelet transform perform a multi-modal split of identity, where orientation information, isotropy information, total amplitude information and (two) phase information are separated for each passband. Hence, the presented approach generalizes the invariance – equivariance property to five independent local signal features on each scale. Although the scale-decomposition is not orthogonal, which means that the signal components at different radial frequencies are not projected onto really independent components, the decomposition is at least approximately orthogonal for frequency components which have a sufficient large distance in Fourier domain.

A couple of applications show that the new framework for image processing simplifies the design of specific algorithms for feature extraction. Although the applications are not the main topic of this rather theoretic thesis, the results of the presented methods are comparable to those of other approaches which are especially designed for a specific purpose. I think that this fact confirms the applicability of the developed theory.

7.2 Further Extensions, Open Problems, and Future Work

The generalization of the analytic signal is restricted to the 2D case in this thesis. However, it is possible the generalize to higher dimensions. In [43] a formulation of the monogenic signal and spherical quadrature filters for arbitrary dimensions is proposed. The monogenic phase in the nD case consists of n angles which represent the local orientation ($n-1$ angles) and the local i1D structure. Other topics which are treated in this thesis like scale-space, structure multivector, and applications have not been generalized to higher dimensions so far. Most of the mathematical framework exists also for higher dimensions and therefore, further generalizations should be possible.

Although most aspects concerning the presented theory are treated in this thesis, there are still some open problems. First of all, the orientation – direction dilemma is not solved totally yet. For signals with an underlying unique phase function, as it is the case for fringe patterns, the problem has been solved in a satisfactory way by considering singularities in the orientation field [93]. For general signals as images

this approach cannot be applied and the problem remains basically open. A second open problem refers to the presented scale-space and the related continuous wavelet transform. For practical applications it is desirable to have a numerical scheme for computing the scale-space from a discretized differential equation. Furthermore, the introduction of the local metric as a regularization would be appropriate. The relation of such an approach to a discrete wavelet transform or fast wavelet transform based on the developed radial bandpass filters is a quite fundamental question. A third open question is the effect of the considered i2D signal model for arbitrary signals. It is worth to consider the error which is introduced by the model and whether it is reasonable to use higher order terms for further improvements of the classification of i2D structures (see also [146] for considerations on this topic). A further quite interesting field which has not been considered in this thesis are vector-valued signals, e.g., color images. The question how to deal with color images in the context of the new quadrature filters remains totally open.

From a practical point of view an adaptive, iterative algorithm based on the mutual estimation of local scale and local isotropy and application of the so adapted filters is desirable. A numerical scheme as it is sketched in the previous paragraph probably behaves in such a way if it is regularized appropriately. The output of this algorithm would consist of stabilized estimates of local (main) orientation, phase(s), and amplitude(s). In my opinion such a set of stabilized estimates is the optimal basis for further processing on the visual data. Since the presented methods are *complete* in a sense that no information is lost (at least in theory), further processing can make use of all the information. On the other hand, the representation of data is changed totally. The new approach fulfills the invariance – equivariance property and hence, further decision-making processing can take advantage of this improved representation since the relevant information is directly accessible. Hence, the output of the structure multivector is probably well suited to be fed into artificial neural networks which will also be topic of future work.

Appendix A

FOURIER TRANSFORM PAIRS AND UNCERTAINTIES

In this appendix some 1D and 2D Fourier transform pairs and the uncertainties of the 1D and 2D Poisson kernels are calculated. Furthermore, a 'one-sided' computation scheme for the monogenic signal is derived.

A.1 1D Fourier Transform Pairs

Throughout this section, 1D variables are denoted as vectors, i.e., $\mathbf{x} = x\mathbf{e}_1$ for the spatial position and $\mathbf{u} = u\mathbf{e}_1$ for the frequency.

Poisson Kernel

The Fourier transform of the 1D Poisson kernel is given as

$$\mathcal{F}_1 \left\{ \frac{y}{\pi |\mathbf{x} + y\mathbf{e}_2|^2} \right\} (\mathbf{u}) = \exp(-2\pi |\mathbf{u}| y)$$

for $y > 0$.

According to the integral table 19 in [16], integral no. 18,

$$\int_0^\infty \frac{\cos(ax)}{1 + x^2}\, dx = \frac{\pi}{2} \exp(-|a|) \ .$$

Applying the substitution $x' = bx$ where $b > 0$ yields

$$\int_0^\infty \frac{b^2 \cos(ab^{-1}x')}{b^2 + x'^2} b^{-1}\, dx' = \frac{\pi}{2} \exp(-|a|) \ .$$

Setting $a' = ab^{-1}$ then gives

$$\int_0^\infty \frac{b}{b^2 + x'^2} \cos(a'x')\, dx' = \frac{\pi}{2} \exp(-|a'b|) \overset{b>0}{=} \frac{\pi}{2} \exp(-|a'|b) \ ,$$

which is the cosine Fourier transform of $\frac{b}{b^2+x^2}$ (a' being the rotation frequency)[1]. Hence, setting $b = y$ and $a' = 2\pi u$

$$\mathcal{F}_1 \left\{ \frac{y}{\pi |\mathbf{x} + y\mathbf{e}_2|^2} \right\} (\mathbf{u}) = \frac{2}{\pi} \int_0^\infty \frac{y}{x^2 + y^2} \cos(2\pi x u)\, dx = \exp(-2\pi |\mathbf{u}| y) \ .$$

Conjugate Poisson Kernel

The Fourier transform of the 1D conjugate Poisson kernel is given as

$$\mathcal{F}_1 \left\{ \frac{x}{\pi |\mathbf{x} + y\mathbf{e}_2|^2} I_2 \right\} (\mathbf{u}) = \frac{u}{|u|} \exp(-2\pi |\mathbf{u}| y)$$

for $y > 0$.

According to the integral table 19 in [16], integral no. 17,

$$\int_0^\infty \frac{x \sin(bx)}{a^2 + x^2}\, dx = \operatorname{sign}(b) \frac{\pi}{2} \exp(-|ab|) \ .$$

Setting $b = 2\pi u$ and $a = y$ the sine Fourier transform of the conjugate Poisson kernel is obtained, and therefore

$$\mathcal{F}_1 \left\{ \frac{x}{\pi |\mathbf{x} + y\mathbf{e}_2|^2} I_2 \right\} (\mathbf{u}) = \operatorname{sign}(u) \exp(-2\pi |uy|) \overset{y>0}{=} \frac{u}{|u|} \exp(-2\pi |\mathbf{u}| y) \ .$$

I have not used the sine Fourier transform table in [16] for a similar reason as in footnote 1.

[1] I do not use the Fourier transform table in [16] since it seems to be inexact. In the Fourier transform table, the cosine Fourier transform of $(a^2 + t^2)^{-1}$ is given without the modulus operator of ω ($\exp(-a\omega)$ instead of $\exp(-a|\omega|)$) and without the restriction $a > 0$. Obviously, this is an error, because the cosine Fourier transform of a real, even function must be even again and its limit for ω tending to infinity must be zero (see (12.67) in the same reference).

A.2 2D Fourier Transform Pairs

Throughout this section, 2D variables are denoted as vectors, i.e., $\mathbf{x} = x\mathbf{e}_1 + y\mathbf{e}_2$ for the spatial position and $\mathbf{u} = u\mathbf{e}_1 + v\mathbf{e}_2$ for the frequency. In some cases it is necessary to use polar coordinates, where $r = \sqrt{\mathbf{x}\mathbf{x}} = |\mathbf{x}|$ and $q = \sqrt{\mathbf{u}\mathbf{u}} = |\mathbf{u}|$.

Riesz Transform

The Fourier transform of the Riesz kernel is given as

$$\mathcal{F}_2\left\{\frac{\mathbf{x}\mathbf{e}_3}{2\pi|\mathbf{x}|^3}\right\}(\mathbf{u}) = \frac{\mathbf{u}}{|\mathbf{u}|}I_2^{-1} \quad .$$

According to the table of Hankel transforms in [12],

$$\mathcal{F}_2\{r^{-1}\}(\mathbf{u}) = q^{-1} \quad .$$

Applying the derivative theorem of the 2D Fourier transform,

$$\mathcal{F}_2\left\{-\frac{\mathbf{x}}{|\mathbf{x}|^3}\right\}(\mathbf{u}) = -2\pi\frac{\mathbf{u}^*}{|\mathbf{u}|}$$

and by multiplying $-\frac{\mathbf{e}_3}{2\pi}$ from the right, the Fourier transform of the Riesz kernel is obtained.

Poisson Kernel

The Fourier transform of the 2D Poisson kernel is given as

$$\mathcal{F}_2\left\{\frac{s}{2\pi|\mathbf{x} + s\mathbf{e}_3|^3}\right\}(\mathbf{u}) = \exp(-2\pi|\mathbf{u}|s)$$

for $s > 0$.

According to the table of Hankel transforms in [12],

$$\mathcal{F}_2\{(1 + r^2)^{-3/2}\}(\mathbf{u}) = 2\pi\exp(-2\pi q) \quad .$$

Substituting $r' = sr$, $s > 0$ and applying the affine theorem yields

$$\mathcal{F}_2\{s^3(s^2 + r'^2)^{-3/2}\}(\mathbf{u}) = s^2 2\pi\exp(-2\pi qs) \quad .$$

By setting $r' = |\mathbf{x}|$ and multiplying with $(s^2 2\pi)^{-1}$ the Fourier transform of the Poisson kernel is obtained.

Conjugate Poisson Kernel

The Fourier transform of the 2D conjugate Poisson kernel is given as

$$\mathcal{F}_2\left\{\frac{\mathbf{x}\mathbf{e}_3}{2\pi|\mathbf{x}+s\mathbf{e}_3|^3}\right\}(\mathbf{u}) = \frac{\mathbf{u}}{|\mathbf{u}|}I_2^{-1}\exp(-2\pi|\mathbf{u}|s)$$

for $s > 0$.

According to the table of Hankel transforms in [12],

$$\mathcal{F}_2\{(1+r^2)^{-1/2}\}(\mathbf{u}) = q^{-1}\exp(-2\pi q)$$

and by substituting $r' = sr$ ($s > 0$)

$$\mathcal{F}_2\{s(s^2+r'^2)^{-1/2}\}(\mathbf{u}) = s^2(qs)^{-1}\exp(-2\pi qs) \ .$$

Hence,

$$\mathcal{F}_2\{(s^2+|\mathbf{x}|^2)^{-1/2}\}(\mathbf{u}) = |\mathbf{u}|^{-1}\exp(-2\pi|\mathbf{u}|s) \ .$$

Applying the derivative theorem yields

$$\mathcal{F}_2\{-\mathbf{x}(s^2+|\mathbf{x}|^2)^{-3/2}\}(\mathbf{u}) = -2\pi\mathbf{u}^*|\mathbf{u}|^{-1}\exp(-2\pi|\mathbf{u}|s)$$

and by multiplying with $-\frac{\mathbf{e}_3}{2\pi}$ from the right the Fourier transform of the conjugate Poisson kernel is obtained.

Damped Spherical Harmonic of Order 2

Considering the conjugate Poisson kernel as a damped spherical harmonic of first order in the frequency domain, allows to define higher order spherical harmonics with the same amplitude spectrum. The Fourier transform of the damped spherical harmonic of order two is given as

$$\mathcal{F}_2\left\{\frac{\mathbf{x}\mathbf{e}_1\mathbf{x}(s(2s^2+3\mathbf{x}^2)-2|\mathbf{x}+s\mathbf{e}_3|^3)}{2\pi\mathbf{x}^4|\mathbf{x}+s\mathbf{e}_3|^3}\right\}(\mathbf{u}) = \frac{\mathbf{u}\mathbf{e}_1\mathbf{u}}{\mathbf{u}^2}\exp(-2\pi|\mathbf{u}|s)$$

for $s > 0$ so that

$$\lim_{s\to 0}\mathcal{F}_2^{-1}\left\{\frac{\mathbf{u}\mathbf{e}_1\mathbf{u}}{\mathbf{u}^2}\exp(-2\pi|\mathbf{u}|s)\right\}(\mathbf{x}) = -\frac{2\mathbf{x}\mathbf{e}_1\mathbf{x}}{2\pi\mathbf{x}^4} \ .$$

According to [136], the 2D Fourier transform of an L^1-integrable function f which can be separated into a radial function and a complex exponential of the angular

coordinate, i.e., $f(x,y) = f_0(r)\exp(ik\phi)$, where $k \in \mathbb{N}$ and $x + iy = r\exp(i\phi)$, is given by

$$2\pi(-i)^k \exp(ik\theta) \int_0^\infty f_0(r) J_k(2\pi qr) r\, dr \tag{A.1}$$

where $u + iv = q\exp(i\theta)$. For the case of $k = 2$ and $f_0(r) = \exp(-2\pi rs)$ with $s > 0$

$$-2\pi \exp(i2\theta) \frac{2(s^2+q^2)^{3/2} - s(2s^2+3q^2)}{4\pi^2 q^2 (s^2+q^2)^{3/2}} = \exp(i2\theta) \frac{s(2s^2+3q^2) - 2(s^2+q^2)^{3/2}}{2\pi q^2 (s^2+q^2)^{3/2}}$$

is obtained (by use of an algebra program). In Cartesian coordinates, this is equivalent to

$$(x+iy)^2 \frac{s(2s^2+3|x+iy|^2) - 2(s^2+|x+iy|^2)^{3/2}}{2\pi |x+iy|^4 (s^2+|x+iy|^2)^{3/2}} \quad .$$

Since both parts (the real part and the imaginary part) are symmetric with respect to the origin, exchanging Fourier and spatial domain does not make any difference. Replacing then i with $\mathbf{e}_{12}$ and multiplying $\mathbf{e}_1$ from the left yields the inverse Fourier transform of $\mathbf{u}\mathbf{e}_1\mathbf{u}|\mathbf{u}|^{-2}\exp(-2\pi|\mathbf{u}|s)$.

Comparing the Riesz transform and the case $s = 0$ of the just derived Fourier transform pair, there seems to be a contradiction to the Parseval theorem since both transfer functions have the same modulus for every frequency vector whereas the modulus of the impulse responses differs by a factor of two for every spatial vector. Actually, the Fourier transform pair is correct and the Parseval theorem is fulfilled for any $s > 0$ which can be checked by calculating the energy in both cases: $E = (8s^2\pi)^{-1}$. However, for $s = 0$ the energy is infinite in both cases and therefore, it is also equal.

On the other hand, taking the limit $s \to 0$ only yields a Fourier transform pair if the Fourier transform of the limit exists. Since I am not interested in applying the undamped second order harmonic (actually, this is not possible at all) this question is only of theoretic relevance. However, according to [12], page 152, the impulse response of the second order harmonic is a Dirac function with an angular second order modulation. This Fourier pair is obviously derived from the Fourier pair just below on the same page, which is a Gaussian with a second order modulation in both domains. The latter Fourier pair is wrong since it contradicts the theorem from [136].

Damped Spherical Harmonic of Order 3

The Fourier transform of the damped spherical harmonic of order three is given as

$$\mathcal{F}_2\left\{\frac{\mathbf{x}\mathbf{e}_1\mathbf{x}\mathbf{e}_1\mathbf{x}\mathbf{e}_3(8s|\mathbf{x}+s\mathbf{e}_3|^3 - 12s^2\mathbf{x}^2 - 8s^4 - 3\mathbf{x}^4)}{2\pi\mathbf{x}^6|\mathbf{x}+s\mathbf{e}_3|^3}\right\}(\mathbf{u}) = \frac{\mathbf{u}\mathbf{e}_1\mathbf{u}\mathbf{e}_1\mathbf{u}}{|\mathbf{u}|^3} I_2^{-1}\exp(-2\pi|\mathbf{u}|s)$$

for $s > 0$ so that

$$\lim_{s\to 0} \mathcal{F}_2^{-1}\left\{\frac{\mathbf{u}\mathbf{e}_1\mathbf{u}\mathbf{e}_1\mathbf{u}}{|\mathbf{u}|^3} I_2^{-1}\exp(-2\pi|\mathbf{u}|s)\right\}(\mathbf{x}) = \frac{-3\mathbf{x}\mathbf{e}_1\mathbf{x}\mathbf{e}_1\mathbf{x}\mathbf{e}_3}{2\pi|\mathbf{x}|^5} \quad .$$

According to (A.1) with $k = 3$

$$i\exp(i3\theta)\frac{12s^2q^2 + 8s^4 + 3q^4 - 8s(s^2+q^2)^{3/2}}{2\pi q^3(s^2+q^2)^{3/2}}$$

is obtained (again using an algebra program). In Cartesian coordinates this is equivalent to

$$i(x+iy)^3\frac{12s^2|x+iy|^2 + 8s^4 + 3|x+iy|^4 - 8s(s^2+|x+iy|^2)^{3/2}}{2\pi|x+iy|^6(s^2+|x+iy|^2)^{3/2}} \quad .$$

Since both parts are antisymmetric with respect to the origin, exchanging Fourier and spatial domain just yields a minus sign. Whereas the leftmost imaginary unit is obtained from the Fourier kernel, i.e., it must be replaced by I_3, the other imaginary units must be replaced with $\mathbf{e}_{12}$. Multiplying both sides with $-\mathbf{e}_2$ from the left yields the inverse Fourier transform of $\mathbf{u}\mathbf{e}_1\mathbf{u}\mathbf{e}_1\mathbf{u}|\mathbf{u}|^{-3}I_2^{-1}\exp(-2\pi|\mathbf{u}|s)$.

Again, the Parseval theorem is fulfilled for $s > 0$ ($E = (8s^2\pi)^{-1}$). The remarks about the limit $s \to 0$ of the second order harmonic apply accordingly.

Radially Weighted Damped Spherical Harmonics

Besides the damped spherical harmonics of higher order, the same filters with an additional factor of $|\mathbf{u}|$ exist. Using (A.1) it can be shown that

$$\mathcal{F}_2\left\{-\frac{3\mathbf{x}\mathbf{e}_3}{4\pi^2|\mathbf{x}+s\mathbf{e}_3|^5}\right\}(\mathbf{u}) = \mathbf{u}I_2^{-1}\exp(-2\pi|\mathbf{u}|s)$$

where $s > 0$. The energy of this LSI operator is $E = \frac{3}{64s^4\pi^3}$. The limit $s \to 0$ does not yield a Fourier transform pair.

Quite similar, it can be shown that

$$\mathcal{F}_2\left\{-\frac{3\mathbf{x}\mathbf{e}_1\mathbf{x}}{4\pi^2|\mathbf{x}+s\mathbf{e}_3|^5}\right\}(\mathbf{u}) = \frac{\mathbf{u}\mathbf{e}_1\mathbf{u}}{|\mathbf{u}|}\exp(-2\pi|\mathbf{u}|s)$$

where $s > 0$. The energy of this LSI operator is also $E = \frac{3}{64s^4\pi^3}$ and the limit $s \to 0$ also does not yield a Fourier transform pair.

Also quite similar, it can be shown that

$$\mathcal{F}_2\left\{\frac{\mathbf{x}\mathbf{e}_1\mathbf{x}\mathbf{e}_1\mathbf{x}\mathbf{e}_3(s(20s^2\mathbf{x}^2 + 8s^4 + 15\mathbf{x}^4) - 8|\mathbf{x}+s\mathbf{e}_3|^5)}{4\pi^2|\mathbf{x}|^6|\mathbf{x}+s\mathbf{e}_3|^5}\right\}(\mathbf{u}) = \frac{\mathbf{u}\mathbf{e}_1\mathbf{u}\mathbf{e}_1\mathbf{u}}{\mathbf{u}^2}I_2^{-1}\exp(-2\pi|\mathbf{u}|s)$$

where $s > 0$. The energy of this LSI operator is also $E = \frac{3}{64s^4\pi^3}$ and the limit $s \to 0$ also does not yield a Fourier transform pair.

A.3 Uncertainties

The uncertainty of an L^2 function f is given by the product of the mean spread in spatial domain and frequency domain. The mean spread in the spatial domain is obtained from the variance of the corresponding probability density function $p(\mathbf{x})$ which is defined by[2]

$$p(\mathbf{x}) = \frac{|f(\mathbf{x})|^2}{\|f\|^2} = \frac{|f(\mathbf{x})|^2}{\int |f(\mathbf{x})|^2 \, d\mathbf{x}}$$

(see e.g. [120]). Hence, the variance is defined by

$$\sigma_{\mathbf{x}}^2 = \int |\mathbf{x}|^2 p(\mathbf{x}) \, d\mathbf{x} = \frac{\int |\mathbf{x}|^2 |f(\mathbf{x})|^2 \, d\mathbf{x}}{\int |f(\mathbf{x})|^2 \, d\mathbf{x}} \ .$$

The spread is obtained by taking the square root:

$$\Delta\mathbf{x} = \sqrt{\sigma_{\mathbf{x}}^2} \ .$$

If F is the Fourier transform of f, the uncertainty of f is given by the product of the spreads in both domains:

$$(\Delta\mathbf{x})(\Delta\mathbf{u}) = \sqrt{\sigma_{\mathbf{x}}^2 \sigma_{\mathbf{u}}^2}$$

where

$$\sigma_{\mathbf{u}}^2 = \frac{\int |\mathbf{u}|^2 |F(\mathbf{u})|^2 \, d\mathbf{u}}{\int |F(\mathbf{u})|^2 \, d\mathbf{u}} \ .$$

Poisson Kernel for 1D Signals

The Poisson kernel for 1D signals reads[3]

$$p_s(x) = (\mathbf{f}_2 \cdot \mathbf{e}_2)(x, s) = \frac{s}{\pi(x^2 + s^2)}$$

and accordingly, its Fourier transform is given by (see section A.1)

$$P_s(u) = \mathcal{F}_1\{p_s\}(u) = \exp(-2\pi|u|s) \ .$$

[2] Note that all integrals in this section are infinite integrals (on the whole domain) except those cases where it is explicitly indicated.

[3] According to section 4.5 the variable s is used instead of y.

The spread in the spatial domain reads

$$\begin{aligned}\Delta x = \sqrt{\sigma_x^2} &= \left(\frac{\int x^2 p_s(x)^2 \, dx}{\int p_s(x)^2 \, dx} \right)^{1/2} \\ &= \left(\frac{\int x^2 \frac{s^2}{\pi^2 (x^2+s^2)^2} \, dx}{\int \frac{s^2}{\pi^2 (x^2+s^2)^2} \, dx} \right)^{1/2} \\ &= \left(\frac{\int \frac{x^2}{(x^2+s^2)^2} \, dx}{\int \frac{1}{(x^2+s^2)^2} \, dx} \right)^{1/2} = \left(\frac{\frac{\pi}{2s}}{\frac{\pi}{2s^3}} \right)^{1/2} = s\end{aligned}$$

where the integrals are evaluated according to [16] 19.5.1.3 integral 58 and 66. The spread in the frequency domain is obtained as

$$\begin{aligned}\Delta u = \sqrt{\sigma_u^2} &= \left(\frac{\int u^2 P_s(u)^2 \, du}{\int P_s(u)^2 \, du} \right)^{1/2} \\ &= \left(\frac{\int u^2 \exp(-4\pi |u| s) \, du}{\int \exp(-4\pi |u| s) \, du} \right)^{1/2} = \left(\frac{\frac{4}{(4\pi s)^3}}{\frac{2}{4\pi s}} \right)^{1/2} = \frac{\sqrt{2}}{4\pi s}\end{aligned}$$

where the integrals are evaluated according to [16] 19.6.1 integral 1. Hence

$$(\Delta x)(\Delta u) = \frac{\sqrt{2}}{4\pi}$$

in the case of the Poisson kernel which means that the uncertainty is slightly worse than for the 1D Gaussian kernel (factor $\sqrt{2}$).

Poisson Kernel for 2D Signals

The Poisson kernel for 2D signals reads

$$p_s(\mathbf{x}) = (\mathbf{f}_3 \cdot \mathbf{e}_3)(\mathbf{x}, s) = \frac{s}{2\pi (\mathbf{x}^2 + s^2)^{\frac{3}{2}}}$$

and accordingly, its Fourier transform is given by (see section A.2)

$$P_s(\mathbf{u}) = \mathcal{F}_2\{p_s\}(\mathbf{u}) = \exp(-2\pi |\mathbf{u}| s) \ .$$

The spread in the spatial domain reads

$$\begin{aligned}\Delta\mathbf{x} = \sqrt{\sigma_{\mathbf{x}}^2} &= \left(\frac{\iint \mathbf{x}^2 p_s(\mathbf{x})^2\,dx\,dy}{\iint p_s(\mathbf{x})^2\,dx\,dy}\right)^{1/2} \\ &= \left(\frac{\iint \mathbf{x}^2 \frac{s^2}{4\pi^2(\mathbf{x}^2+s^2)^3}\,dx\,dy}{\iint \frac{s^2}{4\pi^2(\mathbf{x}^2+s^2)^3}\,dx\,dy}\right)^{1/2} \\ &= \left(\frac{\iint \frac{\mathbf{x}^2}{(\mathbf{x}^2+s^2)^3}\,dx\,dy}{\iint \frac{1}{(\mathbf{x}^2+s^2)^3}\,dx\,dy}\right)^{1/2} \\ &\quad \text{change to polar coordinates } r=|\mathbf{x}|\text{:} \\ &= \left(\frac{2\pi\int_0^\infty \frac{r^3}{(r^2+s^2)^3}\,dr}{2\pi\int_0^\infty \frac{r}{(r^2+s^2)^3}\,dr}\right)^{1/2} = \left(\frac{\frac{1}{2s^2}-\frac{s^2}{4s^4}}{\frac{1}{4s^4}}\right)^{1/2} = s\end{aligned}$$

where the integrals are evaluated according to [16] 19.5.1.3 integral 63 and 71. The spread in the frequency domain is obtained as

$$\begin{aligned}\Delta\mathbf{u} = \sqrt{\sigma_{\mathbf{u}}^2} &= \left(\frac{\iint \mathbf{u}^2 P_s(\mathbf{u})^2\,du\,dv}{\iint P_s(\mathbf{u})^2\,du\,dv}\right)^{1/2} \\ &= \left(\frac{\iint \mathbf{u}^2 \exp(-4\pi|\mathbf{u}|s)\,du\,dv}{\iint \exp(-4\pi|\mathbf{u}|s)\,du\,dv}\right)^{1/2} \\ &\quad \text{change to polar coordinates } q=|\mathbf{u}|\text{:} \\ &= \left(\frac{2\pi\int_0^\infty q^3\exp(-4\pi qs)\,dq}{2\pi\int_o^\infty q\exp(-4\pi qs)\,dq}\right)^{1/2} = \left(\frac{\frac{6}{(4\pi s)^4}}{\frac{1}{(4\pi s)^2}}\right)^{1/2} = \frac{\sqrt{6}}{4\pi s}\end{aligned}$$

where the integrals are evaluated according to [16] 19.6.1 integral 1. Hence

$$(\Delta\mathbf{x})(\Delta\mathbf{u}) = \frac{\sqrt{6}}{4\pi}$$

in the case of the Poisson kernel which means that the uncertainty is slightly worse than for the 2D Gaussian kernel (factor $\sqrt{1.5}$).

Considering the modified uncertainty according to (3.47), the uncertainty of the 2D Poisson kernel splits according to

$$((\Delta\mathbf{x})(\Delta\mathbf{u}))^{\parallel} = \sqrt{3}(4\pi)^{-1} \quad \text{and} \quad ((\Delta\mathbf{x})(\Delta\mathbf{u}))^{\perp} = \sqrt{3}(4\pi)^{-1}\ .$$

This result is obtained by a similar calculation as above with an additional factor $\cos^2\theta$ in the polar integration, which introduces a factor of π instead of 2π.

A.4 Simplified Transfer Function of the Monogenic Signal

For an $\mathbf{e}_3$-valued signal f the monogenic signal $\mathbf{f}_\mathrm{M}$ can be computed in one arbitrary half-space by the simplified formula

$$\mathbf{f}_\mathrm{M}(\mathbf{x}) = \langle 2 \int_0^\infty \int_{\theta_0}^{\theta_0+\pi} (1 + \mathbf{n}_\theta I_2^{-1}) \mathbf{F}(q\mathbf{n}_\theta) \exp(I_3 2\pi q \mathbf{n}_\theta \cdot \mathbf{x}) |q| \, d\theta \, dq \rangle_1$$

This equation is obtained by straightforward calculations:

$$\begin{aligned}
\mathbf{f}_\mathrm{M}(\mathbf{x}) &= \iint \left(1 + \frac{\mathbf{u}}{|\mathbf{u}|} I_2^{-1}\right) \mathbf{F}(\mathbf{u}) \exp(I_3 2\pi \mathbf{u} \cdot \mathbf{x}) \, du \, dv \\
&\quad \text{change to polar coordinates with } \mathbf{u} = q\mathbf{n}_\theta, \text{ where} \\
&\quad \mathbf{n}_\theta = \cos\theta \mathbf{e}_1 + \sin\theta \mathbf{e}_2 \text{ and } q \in \mathbb{R},\ \theta \in [\theta_0, \theta_0 + \pi) \\
&= \int_{-\infty}^{\infty} \int_{\theta_0}^{\theta_0+\pi} (1 + \mathrm{sign}(q) \mathbf{n}_\theta I_2^{-1}) \mathbf{F}(q\mathbf{n}_\theta) \exp(I_3 2\pi q \mathbf{n}_\theta \cdot \mathbf{x}) |q| \, d\theta \, dq \\
&= \int_0^{\infty} \int_{\theta_0}^{\theta_0+\pi} (1 - \mathbf{n}_\theta I_2^{-1}) \mathbf{F}(-q\mathbf{n}_\theta) \exp(-I_3 2\pi q \mathbf{n}_\theta \cdot \mathbf{x}) | -q| \, d\theta \, dq \\
&\quad + \int_0^{\infty} \int_{\theta_0}^{\theta_0+\pi} (1 + \mathbf{n}_\theta I_2^{-1}) \mathbf{F}(q\mathbf{n}_\theta) \exp(I_3 2\pi q \mathbf{n}_\theta \cdot \mathbf{x}) |q| \, d\theta \, dq \\
&\quad \text{apply the symmetry of the 2D Fourier transform } \mathbf{F}(-\mathbf{u}) = \tilde{\mathbf{F}}(\mathbf{u}) \\
&= \int_0^{\infty} \int_{\theta_0}^{\theta_0+\pi} (1 - \mathbf{n}_\theta I_2^{-1}) \tilde{\mathbf{F}}(q\mathbf{n}_\theta) (\exp(I_3 2\pi q \mathbf{n}_\theta \cdot \mathbf{x}))^{\sim} |q| \, d\theta \, dq \\
&\quad + \int_0^{\infty} \int_{\theta_0}^{\theta_0+\pi} (1 + \mathbf{n}_\theta I_2^{-1}) \mathbf{F}(q\mathbf{n}_\theta) \exp(I_3 2\pi q \mathbf{n}_\theta \cdot \mathbf{x}) |q| \, d\theta \, dq \\
&\quad \text{according to (2.31) this can be reduced to} \\
&= 2 \int_0^{\infty} \int_{\theta_0}^{\theta_0+\pi} (\langle \mathbf{F}(q\mathbf{n}_\theta) \exp(I_3 2\pi q \mathbf{n}_\theta \cdot \mathbf{x}) \rangle_1 \\
&\qquad\qquad + \mathbf{n}_\theta I_2^{-1} \langle \mathbf{F}(q\mathbf{n}_\theta) \exp(I_3 2\pi q \mathbf{n}_\theta \cdot \mathbf{x}) \rangle_2) |q| \, d\theta \, dq \\
&\quad \text{and since } \mathbf{F} \text{ only contains } \mathbf{e}_3 \text{ and } \mathbf{e}_{12} \text{ components,} \\
&= 2 \int_0^{\infty} \int_{\theta_0}^{\theta_0+\pi} \langle (1 + \mathbf{n}_\theta I_2^{-1}) \mathbf{F}(q\mathbf{n}_\theta) \exp(I_3 2\pi q \mathbf{n}_\theta \cdot \mathbf{x}) \rangle_1 |q| \, d\theta \, dq \quad ,
\end{aligned}$$

so that the simplified formula for $\mathbf{f}_\mathrm{M}$ follows by linearity.

BIBLIOGRAPHY

[1] ABLAMOWICZ, R. *Clifford Algebras with Numeric and Symbolic Computations*. Birkhäuser Bosten, 1996.

[2] ANDERSSON, M. T. *Controllable Multidimensional Filters and Models in Low Level Computer Vision*. PhD thesis, Linköping University, September 1992.

[3] BABAUD, J., WITKIN, A. P., BAUDIN, M., AND DUDA, R. O. Uniqueness of the gaussian kernel for scale-space filtering. *IEEE Transactions on Pattern Analysis and Machine Intelligence 8*, 1 (1986), 26–33.

[4] BÅRMAN, H. *Hierarchical Curvature Estimation in Computer Vision*. PhD thesis, Linköping University, September 1991.

[5] BÅRMAN, H., GRANLUND, G. H., AND KNUTSSON, H. A new approach to curvature estimation and description. In *3rd International Conference on Image Processing and its applications* (Warwick, Great Britain, 1989), IEE, pp. 54–58.

[6] BAYLIS, W. E. *Clifford (Geometric) Algebras with Applications in Physics, Mathematics and Engineering*. Birkhäuser, Boston, 1996.

[7] BIGÜN, J., AND GRANLUND, G. H. Optimal orientation detection of linear symmetry. In *Proceedings of the IEEE First International Conference on Computer Vision* (London, Great Britain, June 1987), pp. 433–438. Report LiTH-ISY-I-0828, Computer Vision Laboratory, Linköping University, Sweden, 1986.

[8] BIGÜN, J., GRANLUND, G. H., AND WIKLUND, J. Multidimensional orientation estimation with applications to texture analysis and optical flow. *IEEE Transactions on Pattern Analysis and Machine Intelligence 13*, 8 (August 1991), 775–790.

[9] BISHOP, C. M. *Neural Networks for Pattern Recognition*. Oxford University Press, New York, 1995.

[10] BOSE, N. K. Trends in multidimensional systems theory. In *Multidimensional Systems Theory*, N. K. Bose, Ed., D. Reidel Publishing, Dordrecht, 1985, pp. 1–40.

[11] BRACEWELL, R. N. *The Fourier transform and its applications*. McGraw Hill, 1986.

[12] BRACEWELL, R. N. *Two-Dimensional Imaging*. Prentice Hall Signal Processing Series. Prentice Hall, Englewood Cliffs, 1995.

[13] BRACKX, F., DELANGHE, R., AND SOMMEN, F. *Clifford Analysis*. Pitman, Boston, 1982.

[14] BRACKX, F., AND SOMMEN, F. Clifford-hermite wavelets in Euclidean space. *The Journal of Fourier Analysis and Applications 6*, 3 (2000), 299–310.

[15] BRADY, J. M., AND HORN, B. M. P. Rotationally symmetric operators for surface interpolation. *Computer Vision, Graphics, and Image Processing 22*, 1 (April 1983), 70–94.

[16] BRONSTEIN, I., SEMENDJAJEW, K., MUSIOL, G., AND MÜHLIG, H. *Taschenbuch der Mathematik*. Verlag Harri Deutsch, Frankfurt, 1993.

[17] BÜLOW, T. *Hypercomplex Spectral Signal Representations for the Processing and Analysis of Images*. PhD thesis, Christian-Albrechts-University of Kiel, 1999.

[18] BÜLOW, T. Consistent phase estimation for recovering depth from moiré interferograms. Tech. rep., Department of Computer and Information Science, University of Pennsylvania, 2001.

[19] BÜLOW, T., FELSBERG, M., AND SOMMER, G. Non-commutative hypercomplex Fourier transforms of multidimensional signals. In *Geometric Computing with Clifford Algebras*, G. Sommer, Ed., Springer-Verlag, Heidelberg, 2001.

[20] BÜLOW, T., PALLEK, D., AND SOMMER, G. Riesz transforms for the isotropic estimation of the local phase of moire interferograms. In *22. DAGM Symposium Mustererkennung, Kiel* (2000), G. Sommer, N. Krüger, and C. Perwass, Eds., Springer-Verlag, Heidelberg.

[21] BÜLOW, T., AND SOMMER, G. Algebraically extended representation of multidimensional signals. In *Proceedings of the 10th Scandinavian Conference on Image Analysis* (1997), pp. 559–566.

[22] BÜLOW, T., AND SOMMER, G. Multi-dimensional signal processing using an algebraically extended signal representation. In *Int'l Workshop on Algebraic Frames for the Perception-Action Cycle, AFPAC'97, Kiel* (1997), G. Sommer and J. J. Koenderink, Eds., vol. 1315 of *LNCS*, Springer-Verlag, Berlin, pp. 148–163.

[23] BÜLOW, T., AND SOMMER, G. The hypercomplex signal – a novel approach to the multidimensional analytic signal. *IEEE Transactions on Signal Processing 49*, 11 (2001), 2844–2852.

[24] BURG, K., HAF, H., AND WILLE, F. *Höhere Mathematik für Ingenieure, Band II Lineare Algebra*. Teubner Stuttgart, 1992.

[25] BURG, K., HAF, H., AND WILLE, F. *Höhere Mathematik für Ingenieure, Band V Funktionalanalysis und Partielle Differentialgleichungen*. Teubner Stuttgart, 1993.

[26] BURG, K., HAF, H., AND WILLE, F. *Höhere Mathematik für Ingenieure, Band IV Vektoranalysis und Funktionentheorie*. Teubner Stuttgart, 1994.

[27] BURT, P. J., AND ADELSON, E. H. The Laplacian pyramid as a compact image code. *IEEE Trans. Communications 31*, 4 (1983), 532–540.

[28] CANNY, J. A computational approach to edge detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence 8*, 6 (November 1986), 679–698.

[29] CATTÉ, F., LIONS, P.-L., MOREL, J.-M., AND COLL, T. Image selective smoothing and edge detection by nonlinear diffusion. *SIAM J. Numer. Analysis 32* (1992), 1895–1909.

[30] CHEN, Q., DEFRISE, M., AND DECONINCK, F. Symmetric phase-only matched filtering of Fourier-Mellin transforms for image registration and recognition. *Transactions on Pattern Analysis and Machine Intelligence 16*, 12 (December 1994), 1156–1168.

[31] DAVENPORT, C. M. A commutative hypercomplex algebra with associated function theory. In *Clifford Algebras with Numeric and Symbolic Computations*, R. Ablamowicz, Ed., Birkhäuser Bosten, 1996, pp. 213–227.

[32] DAVIS, J. A., MCNAMARA, D. E., COTTRELL, D. M., AND CAMPOS, J. Image processing with the radial Hilbert transform: theory and experiments. *Optics Letters 25*, 2 (January 2000), 99–101.

[33] DELANGHE, R., SOMMEN, F., AND SOUČEK, V. *Clifford Algebra and Spinor-Valued Functions*. Kluwer Academic Publishers, Dordrecht, 1992.

[34] DORAN, C. J. L., HESTENES, D., SOMMEN, F., AND VAN ACKER, N. Lie groups as spin groups. *J. Math. Phys. 34*, 8 (August 1993), 3642–3669.

[35] DORAN, C. J. L., AND LASENBY, A. N. Physical applications of geometric algebra. `http://www.mrao.cam.ac.uk/~clifford/ptIIIcourse/`, 2001. (Accessed 19 Nov 2001).

[36] DORST, L., MANN, S., AND BOUMA, T. GABLE: A Matlab geometric algebra tutorial. `http://carol.wins.uva.nl/~leo/clifford/gable.html`, August 2000. (Accessed 19 Nov 2001).

[37] ELL, T. A. *Hypercomplex Spectral Transformations*. PhD thesis, University of Minnesota, 1992.

[38] ERNST, R. R., AUE, W. P., BACHMANN, P., KARHAN, J., KUMAR, A., AND MÜLLER, L. Two-dimensional NMR spectroscopy. In *Proc. 4th Ampère Int. Summer School, Pula, Yugoslavia* (1976).

[39] FARNEBÄCK, G. Spatial domain methods for orientation and velocity estimation. Lic. Thesis LiU-Tek-Lic-1999:13, Dept. EE, Linköping University, March 1999.

[40] FELSBERG, M. Signal processing using frequency domain methods in Clifford algebra. Diploma thesis, Institute of Computer Science and Applied Mathematics, Christian-Albrechts-University of Kiel, 1998.

[41] FELSBERG, M. Homepage. `http://www.ks.informatik.uni-kiel.de/~mfe`, November 2001. (Accessed 19 Nov 2001).

[42] FELSBERG, M., BÜLOW, T., AND SOMMER, G. Commutative hypercomplex Fourier transforms of multidimensional signals. In *Geometric Computing with Clifford Algebras*, G. Sommer, Ed., Springer-Verlag, Heidelberg, 2001.

[43] FELSBERG, M., AND SOMMER, G. The multidimensional isotropic generalization of quadrature filters in geometric algebra. In *Proc. Int. Workshop on Algebraic Frames for the Perception-Action Cycle* (Kiel, Germany, September 2000), S. G. and Y. Y. Zeevi, Eds., vol. 1888 of *Lecture Notes in Computer Science*, Springer-Verlag, Heidelberg, pp. 175–185.

[44] FELSBERG, M., AND SOMMER, G. A new extension of linear signal processing for estimating local properties and detecting features. In *22. DAGM Symposium Mustererkennung, Kiel* (2000), G. Sommer, N. Krüger, and C. Perwass, Eds., Springer-Verlag, Heidelberg, pp. 195–202.

[45] FELSBERG, M., AND SOMMER, G. Structure multivector for local analysis of images. Tech. Rep. 2001, Institute of Computer Science and Applied Mathematics, Christian-Albrechts-University of Kiel, Germany, February 2000. available at `http://www.ks.informatik.uni-kiel.de`.

[46] FELSBERG, M., AND SOMMER, G. The monogenic signal. *IEEE Transactions on Signal Processing 49*, 12 (December 2001), 3136–3144.

[47] FELSBERG, M., AND SOMMER, G. The monogenic signal. Tech. Rep. 2016, Institute of Computer Science and Applied Mathematics, Christian-Albrechts-University of Kiel, Germany, May 2001.

[48] FELSBERG, M., AND SOMMER, G. Scale adaptive filtering derived from the Laplace equation. In *23. DAGM Symposium Mustererkennung, München* (2001), B. Radig and S. Florczyk, Eds., vol. 2191 of *Lecture Notes in Computer Science*, Springer-Verlag, Heidelberg, pp. 124–131.

[49] FELSBERG, M., AND SOMMER, G. The structure multivector. In *Applied Geometrical Algebras in Computer Science and Engineering* (2001), Birkhäuser. to be published.

[50] FELSBERG, M., AND SOMMER, G. Structure multivector for local analysis of images. In *Multi-Image Analysis*, R. Klette, T. Huang, and G. Gimel'farb, Eds., vol. 2032 of *Lecture Notes in Computer Science*. Springer-Verlag, Berlin, 2001, pp. 95–106.

[51] FERRARO, M., AND CAELLI, T. M. Lie transformation groups, integral transforms, and invariant pattern recognition. *Spatial Vision 8*, 4 (1994), 33–44.

[52] FORSSÉN, P.-E. Sparse representations for medium level vision. Lic. Thesis LiU-Tek-Lic-2001:, Dept. EE, Linköping University, February 2001.

[53] FORSSÉN, P.-E., AND GRANLUND, G. Sparse feature maps in a scale hierarchy. In *Proc. Int. Workshop on Algebraic Frames for the Perception-Action Cycle* (Kiel, Germany, September 2000), G. Sommer and Y. Y. Zeevi, Eds., vol. 1888 of *Lecture Notes in Computer Science*, Springer-Verlag, Heidelberg.

[54] FÖRSTNER, W. *Statistische Verfahren für die automatische Bildanalyse und ihre Bewertung bei der Objekterkennung und -vermessung*. No. 370 in C. Verlag der Bayerischen Akademie der Wissenschaften, 1991.

[55] FÖRSTNER, W., AND GÜLCH, E. A fast operator for detection and precise location of distinct points, corners and centres of circular features. In *ISPRS Intercommission Workshop, Interlaken* (June 1987), pp. 149–155.

[56] FRANKEL, T. *The Geometry of Physic*. Cambridge University Press, 1997.

[57] FREEMAN, W. T., AND ADELSON, E. H. The design and use of steerable filters. *IEEE Transactions on Pattern Analysis and Machine Intelligence 13*, 9 (September 1991), 891–906.

[58] GILBERT, J. E., AND MURRAY, M. A. M. *Clifford algebras and Dirac operators in harmonic analysis*. Cambridge University Press, 1991.

[59] GRANLUND, G. H. Hierarchical computer vision. In *Proceedings of EUSIPCO 90, Fifth European Signal Processing Conference* (Barcelona, Spain, September 1990), L. Torres, E. Masgrau, and M. A. Lagunas, Eds., pp. 73–84.

[60] GRANLUND, G. H. The complexity of vision. *Signal Processing 74* (1999), 101–126.

[61] GRANLUND, G. H., AND KNUTSSON, H. *Signal Processing for Computer Vision*. Kluwer Academic Publishers, Dordrecht, 1995.

[62] GRAY, A. *Modern Differential Geometry of Curves and Surfaces*. CRC Press, Boca Raton, 1993.

[63] GULL, S. F., LASENBY, A. N., AND DORAN, C. J. L. Imaginary numbers are not real - the geometric algebra of space time. *Found. Phys. 23*, 9 (1993), 1175.

[64] HAGLUND, L. *Adaptive Multidimensional Filtering*. PhD thesis, Linköping University, October 1992.

[65] HAHN, S. L. Multidimensional complex signals with single-orthant spectra. *Proc. IEEE 80*, 8 (1992), 1287–1300.

[66] HAHN, S. L. *Hilbert Transforms in Signal Processing*. Artech House, Boston, London, 1996.

[67] HANSEN, M. *Stereosehen - ein verhaltensbasierter Ansatz*. PhD thesis, Inst. f. Inf. u. Prakt. Math. der Christian-Albrechts-Universität Kiel, 1998.

[68] HARRIS, C. G., AND STEPHENS, M. A combined corner and edge detector. In *4th Alvey Vision Conference* (1988), pp. 147–151.

[69] HEIN, W. *Struktur- und Darstellungstheorie der klassischen Gruppen*. Springer-Verlag, Berlin, 1990.

[70] HESTENES, D. *New Foundations for Classical Mechanics*. D. Reidel Publishing, Dordrecht, 1986.

[71] HESTENES, D., AND SOBCZYK, G. *Clifford Algebra to Geometric Calculus, A Unified Language for Mathematics and Physics*. Reidel, Dordrecht, 1984.

[72] HORNER, J. L., AND GIANINO, P. D. Phase-only matched filtering. *Applied Optics 23*, 6 (15 March 1984), 812–816.

[73] IIJIMA, T. Basic theory of pattern observation. In *Papers of Technical Group on Automata and Automatic Control, IECE, Japan* (December 1959).

[74] JÄHNE, B. *Digitale Bildverarbeitung*. Springer-Verlag, Berlin, 1997.

[75] JÄHNE, B., HAUSSECKER, J., AND GEISSLER, P. *Handbook of Computer Vision and Applications*. Academic Press, Boston, 1999.

[76] JAIN, R., KASTURI, R., AND SCHUNCK, B. G. *Machine Vision*. McGraw-Hill, New York, 1995.

[77] JONES, J. P., AND PALMER, L. A. An evaluation of the two dimensional gabor filter model of simple receptive fields in striate cortex. *Journal of Neurophysiology 58*, 6 (1987), 1223–1258.

[78] KANTOR, I. L., AND SOLODOVNIKOV, A. S. *Hypercomplex Numbers*. Springer-Verlag, New-York, 1989.

[79] KNUTSSON, H. Edupack: Orientation. `http://www.isy.liu.se/cvl/Education/Edupack/index.html`, March 2001. (Accessed 06 Mar 2002).

[80] KOENDERINK, J. J. What is a "feature"? *Journal of Intelligent Systems 3*, 1 (1993), 49–82.

[81] KOENDERINK, J. J. A generic framework for image geometry. In *Applied Geometrical Algebras in Computer Science and Engineering* (2001), Birkhäuser. to be published.

[82] KOVESI, P. *Invariant Measures of Image Features from Phase Information*. PhD thesis, University of Western Australia, 1996.

[83] KOVESI, P. Image features from phase information. *Videre: Journal of Computer Vision Research 1*, 3 (1999).

[84] KRABS, W. *Mathematical Foundations of Signal Theory*. Heldermann Verlag, Berlin, 1995.

[85] KRANTZ, S. G. *Function Theory of Several Complex Variables*. Wadsworth & Brooks/Cole, Pacific Grove, 1992.

[86] KRANTZ, S. G. *Handbook of Complex Variables*. Birkhäuser, Boston, 1999.

[87] KRIEGER, G. *Nichtlineare Informationsverarbeitung in biologischen und technischen Sehsystemen: eine Analyse mit Volterra-Reihen und Statistiken höherer Ordnung*. PhD thesis, Technische Universität München, 2000.

[88] KRIEGER, G., AND ZETZSCHE, C. Nonlinear image operators for the evaluation of local intrinsic dimensionality. *IEEE Transactions on Image Processing 5*, 6 (June 1996), 1026–1041.

[89] KRÜGER, N., ACKERMANN, M., AND SOMMER, G. Accumulation of 3D object representations within a perception-action-cycle. In *2nd Int. ICSC Symposium on Engineering of Intelligent Systems (EIS'2000), University of Paisley, Scotland* (2000), pp. 116–122.

[90] KRÜGER, N., ACKERMANN, M., AND SOMMER, G. Accumulation of object representations utilizing interaction of robot action and perception. In *22. DAGM Symposium Mustererkennung, Kiel* (2000), G. Sommer, N. Krüger, and C. Perwass, Eds., Springer-Verlag, pp. 365–372.

[91] LAPLANTE, P. A., Ed. *Electrical Engineering Dictionary*. CRC Press, Boca Raton, 2000.

[92] LARKIN, K. G. Natural demodulation of two-dimensional fringe patterns: II. stationary phase analysis of the spiral phase quadrature transform. *Journal of the Optical Society of America A 18*, 8 (2001), 1871–1881.

[93] LARKIN, K. G., BONE, D. J., AND OLDFIELD, M. A. Natural demodulation of two-dimensional fringe patterns: I. general background of the spiral phase quadrature transform. *Journal of the Optical Society of America A 18*, 8 (2001), 1862–1870.

[94] LI, C., MCINTOSH, A., AND QIAN, T. Clifford algebras, Fourier transforms, and singular convolution operators on Lipschitz surfaces. *Revista Matemática Iberoamericana 10*, 3 (1994), 665–721.

[95] LINDEBERG, T. *Scale-Space Theory in Computer Vision*. The Kluwer International Series in Engineering and Computer Science. Kluwer Academic Publishers, Boston, 1994.

[96] LOUIS, A. K., MAASS, P., AND RIEDER, A. *Wavelets*. Teubner, Stuttgart, 1994.

[97] LOUNESTO, P. *Clifford Algebras and Spinors*, vol. 239 of *London Mathematical Society Lecture Note Series*. Cambridge University Press, 1997.

[98] MCINTOSH, A. Clifford algebras, Fourier theory, singular integrals, and harmonic functions on Lipschitz domains. In *Clifford Algebras in Analysis and Related Topics*, J. Ryan, Ed., Studies in Advanced Mathematics. CRC Press, Boca Raton, 1996, pp. 33–88.

[99] MERRON, J., AND BRADY, M. Isotropic gradient estimation. In *IEEE Computer Vision and Pattern Recognition* (1996), pp. 652–659.

[100] MEYER, Y., AND COIFMAN, R. *Wavelets: Calderón-Zygmund and multilinear operators*. Cambridge University Press, 1997.

[101] MICHAELIS, M. *Low Level Image Processing Using Steerable Filters*. PhD thesis, Christian-Albrechts-University of Kiel, 1995.

[102] MICHAELIS, M., AND SOMMER, G. A Lie group approach to steerable filters. *Pattern Recognition Letters 16* (1995), 1165–1174.

[103] MILKEREIT, B., AND SPENCER, C. Multiattribute processing of seismic data: Application to dip displays. *Canadian Journal of Exploration Geophysics 26*, 1 & 2 (December 1990), 47–53.

[104] MITREA, M. Hypercomplex variable techniques in harmonic analysis. In *Clifford Algebras in Analysis and Related Topics*, J. Ryan, Ed., CRC Press, Boca Raton, 1996, pp. 103–128.

[105] MOKHTARIAN, F. Fm: Image corner detection through curvature scale space. `http://www.ee.surrey.ac.uk/Research/VSSP/demos/corners/`, June 2001. (Accessed 16 Nov 2001).

[106] MOKHTARIAN, F., AND SUOMELA, R. Robust image corner detection through curvature scale space. *IEEE Trans. Pattern Analysis and Machine Intelligence 20*, 12 (1998), 1376–1381.

[107] MORAN, B. Mathematics of radar. In *Proceedings of the 20th Century Harmonic Analysis conference* (2001), Kluwer.

[108] MÜLLER, M., AND MARQUARD, J. Die Hilberttransformation und ihre Verallgemeinerung in Optik und Bildverarbeitung. *Optik 110*, 2 (1999), 99–109.

[109] NABIGHIAN, M. N. Toward a three-dimensional automatic interpretation of potential field data via generalized Hilbert transforms: Fundamental relations. *Geophysics 49*, 6 (June 1984), 780–786.

[110] NIETO-VESPERINAS, M. Dispersion relations in two dimensions: Application to the phase problem. *Optik 56*, 4 (1980), 377–384.

[111] OPPENHEIM, A., AND LIM, J. The importance of phase in signals. *Proc. of the IEEE 69*, 5 (May 1981), 529–541.

[112] OSTEN, W. *Digitale Verarbeitung und Auswertung von Interferenzbildern*. Akademie Verlag, Berlin, 1991.

[113] PAPOULIS, A. *Probability, Random Variables and Stochastic Processes*. McGraw-Hill, 1965.

[114] PAPOULIS, A. *Signal Analysis*. McGraw-Hill, 1977.

[115] PENNEC, X., AND AYACHE, N. Uniform distribution, distance and expectation problems for geometric features processing. *Journal of Mathematical Imaging and Vision*, 9 (1998), 49–67.

[116] PERONA, P., AND MALIK, J. Scale-space and edge detection using anisotropic diffusion. *IEEE Trans. Pattern Analysis and Machine Intelligence 12*, 7 (1990), 629–639.

[117] PÖRKSEN, M. 3D Objektrepräsentation basierend auf dem Strukturmultivektor. Diploma thesis, Institut für Informatik und Praktische Mathematik der Christian-Albrechts-Universität zu Kiel, 2001.

[118] PORTEOUS, I. R. *Clifford Algebras and the Classical Groups*. Cambridge University Press, 1995.

[119] RADON, J. On the determination of functions from their integral values along certain manifolds. *IEEE Transactions on Medical Imaging 5*, 4 (December 1986), 170–176. Translation of the original German text by P.C.Parks.

[120] RAMANATHAN, J. *Methods of Applied Fourier Analysis*. Birkhäuser, 1998.

[121] REISFELD, D. The Constrained Phase Congruency feature detector: Simultaneous localization, classification and scale determination. *Pattern Recognition Letters 17* (1996), 1161–1169.

[122] RUBINSTEIN, J., SEGMAN, J., AND ZEEVI, Y. Y. Recognition of distorted patterns by invariance kernels. *Pattern Recognition 24*, 10 (1991), 959–967.

[123] RYAN, J., AND SPRÖSSIG, W., Eds. *Clifford Algebras and their Applications in Mathematical Physics, Volume 2, Clifford Analysis*, vol. 19 of *Progress in Physics*. Birkhäuser, Boston, 2000.

[124] SANGWINE, S. J. Fourier-transforms of color images using quaternion or hypercomplex numbers. *Electronic Letters 32*, 21 (October 1996), 1979–1980.

[125] SATTINGER, D. H., AND WEAVER, O. L. *Lie Groups and Algebras with Applications to Physics, Geometry, and Mechanics*. Springer-Verlag, New York, 1993.

[126] SCHARR, H., KÖRKEL, S., AND JÄHNE, B. Numerische Isotropieoptimierung von FIR-Filtern mittels Querglättung. In *19. DAGM Symposium Mustererkennung, Braunschweig* (1997), E.Paulus and F.M.Wahl, Eds., pp. 367–374.

[127] SCHIFF, J. L. *The Laplace Transform*. Undergraduate Texts in Mathematics. Springer-Verlag, New York, 1999.

[128] SCHRÖDER, H. *Mehrdimensionale Signalverarbeitung*. Teubner, Stuttgart, 1998.

[129] SCHÜSSLER, H. W. *Digitale Signalverarbeitung 1*. Springer-Verlag, Berlin, 1994.

[130] SEGMAN, J., RUBINSTEIN, J., AND ZEEVI, Y. Y. The canonical coordinates method for pattern deformation: Theoretical and computational considerations. *IEEE Transactions on Pattern Analysis and Machine Intelligence 14*, 12 (December 1992), 1171–1183.

[131] SEGMAN, J., AND ZEEVI, Y. Y. Image analysis by wavelet-type transforms: Group theoretic approach. *Journal of Mathematical Imaging and Vision 3* (1993), 51–77.

[132] SIMONCELLI, E. P., AND FARID, H. Steerable wedge filters for local orientation analysis. *IEEE Transactions on Image Processing 5*, 9 (1996), 1377–1382.

[133] SMITH III, J. O. Frequency response. `http://www-ccrma.stanford.edu/~jos/filters/Frequency_Response.html`, September 2001. (Accessed 3 Dec 2001).

[134] SOCHEN, N., KIMMEL, R., AND MALLADI, R. A geometrical framework for low level vision. *IEEE Trans. on Image Processing, Special Issue on PDE based Image Processing 7*, 3 (1998), 310–318.

[135] SOMMER, G. The global algebraic frame of the perception-action cycle. In *Handbook of Computer Vision and Applications* (1999), B. Jähne, H. Haußecker, and P. Geissler, Eds., vol. 3, Academic Press, San Diego, pp. 221–264.

[136] STEIN, E., AND WEISS, G. *Introduction to Fourier Analysis on Euclidean Spaces*. Princeton University Press, New Jersey, 1971.

[137] STEIN, E. M. *Singular Integrals and Differentiability Properties of Functions*. Princeton University Press, New Jersey, 1970.

[138] STEIN, E. M. *Harmonic Analysis*. Princeton University Press, New Jersey, 1993.

[139] STEIN, E. M. Singular integrals: The roles of Calderón and Zygmund. *Notices of the American Mathematical Society 45*, 9 (1998), 1130–1140.

[140] TAO, T. Lecture about harmonic analysis in the phase plane: Review of fourier transform, phase space, uncertainty principle. `http://www.math.ucla.edu/~tao/254a.1.01w/notes1.dvi`, January 2001. (Accessed 19 Nov 2001).

[141] VAN GOOL, L., MOONS, T., PAUWELS, E., AND OOSTERLINCK, A. Vision and Lie's approach to invariance. *Image and Vision Computing 13*, 4 (May 1995), 259–277.

[142] WEICKERT, J. *Anisotropic Diffusion in Image Processing*. PhD thesis, Faculty of Mathematics, University of Kaiserslautern, 1996.

[143] WEICKERT, J. A review of nonlinear diffusion filtering. In *Scale-Space Theory in Computer Vision* (1997), B. ter Haar Romeny, L. Florack, J. Koenderink, and M. Viergever, Eds., vol. 1252 of *LNCS*, Springer, Berlin, pp. 260–271.

[144] WEICKERT, J., ISHIKAWA, S., AND IMIYA, A. Scale-space has first been proposed in Japan. *Mathematical Imaging and Vision 10* (1999), 237–252.

[145] WOODARD, M. R. Quotes from the mathematical quotations server. `http://math.furman.edu/mqs.html`, September 2000. (Accessed 19 Nov 2001).

[146] YU, W., DANIILIDIS, K., BEAUCHEMIN, S., AND SOMMER, G. Detection and characterization of multiple motion points. In *18th IEEE Conference on Computer Vision and Pattern Recognition, Fort Collins, Colorado* (1999).

[147] ZAYED, A. I. *Handbook of Function and Generalized Function Transformations*. Mathematical Science Reference Series. CRC Press, Boca Raton, 1996.

INDEX