Technische Universität München
Fakultät für Mathematik

Monogenic Wavelet Frames for Image Analysis

Stefan Held

Vollständiger Abdruck der von der Fakultät für Mathematik der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften (Dr. rer. nat.)

genehmigten Dissertation.

Vorsitzender:		Univ.-Prof. Dr. Martin Brokate
Prüfer der Dissertation:	1.	Univ.-Prof. Dr. Brigitte Forster-Heinlein
	2.	Univ.-Prof. Dr. Uwe Kähler, Universitário de Santiago / Portugal
	3.	Univ.-Prof. Dr. Rupert Lasser

Die Dissertation wurde am 12. Juli 2012 bei der Technischen Universität München eingereicht
und durch die Fakultät für Mathematik am 22.November 2012 angenommen.

Abstract

The subject of this thesis is the implementation of the Riesz transform via wavelet frames. This leads to an extension of the analytical wavelets, called monogenic wavelets, which allow to decompose an image into amplitude, phase, and a phase direction.

The reasons to use Riesz transforms for image analysis are the phase amplitude decomposition they achieve via monogenic wavelet frames and the steerability of the resulting filters. Wavelets are the best choice for the implementation because they yield an amplitude phase decomposition localized with respect to space and scale. Furthermore, wavelets are especially well adapted to Riesz transforms, since their generating operators commute with Riesz transforms.

We start with an investigation of the Riesz transform and present a novel definition of Riesz transforms of distributions, which results in an extension of the Bedrosian identity.

This is followed by a chapter on Clifford algebras – the language in which many of the following results are formulated. After a short introduction on Clifford algebra we introduce Clifford-Hilbert modules to state new results on left linear operators on Clifford-Hilbert modules.

These results are then used to define our Riesz transform as a steerable, left-linear, self-adjoint, unitary operator on a Clifford-Hilbert module. From this operator we define the monogenic signal as extension of the analytical signal to higher dimensions. The monogenic signal yields a decomposition of a signal into amplitude, phase, and a phase direction, which gives the preferred direction of the signal.

Using a new frame concept on Clifford-Hilbert modules – Clifford frames – we proof that due to its unitarity the Riesz transform maps frames onto Clifford frames. As a last step towards our goal we find suitable (i.e. radial) wavelet frames and implement them using the fast Fourier transform.

The result is a steerable multiwavelet frame in arbitrary dimensions that yields a decomposition into amplitude, phase and a direction which is local with respect to location and scale. Two applications are presented to show the usefulness of the constructed wavelets.

Then we study two tools, which can be used to improve filter design. The first one is the influence of the Riesz transform on the Weyl-Heisenberg uncertainty relation. This allows to study the joint space-time localization of filters and their Riesz transforms. The second one is a sufficient condition for the Gibbs phenomenon to occur for wavelet frames.

The operator resulting from successively applying several partial Riesz transforms is called a higher Riesz transform. We formulate higher Riesz transforms which are steerable unitary left linear operators on Clifford-Hilbert modules, and use the theory we developed earlier for the Riesz transform to implement the higher Riesz transforms. The higher Riesz transforms give rise to a higher monogenic signal, which yields a decomposition into amplitude, phase, and geometrical information. Finally we explain how the steerable filters devised by Freeman and Adelson in [18] can be viewed as higher Riesz transforms applied to radial filters.

Für meine Frau Anja, und meine Töchter Freya und Merle.

Mein Dank gilt:

Prof. Dr. Brigitte Forster, für Ihre Unterstützung und Förderung.

Prof. Dr. Rupert Lasser, der mit seinen Vorlesungen mein Verständniss von der Analysis geprägt hat.

Prof. Dr. Brokate für seine Gastfreundschaft am Lehrstuhl M6.

Peter Massopust der stets ein offenes Ohr für Fragen hatte und fürs Lesen der Arbeit.

Contents

Chapter 1

Introduction

1.1 Introduction

This thesis centers around two ideas that are important in signal and image processing.

Phase information is a valuable tool in signal processing, where it is acquired using analytical wavelets. (For details on analytical wavelets see [43] and the references therein.) Phase information is desirable for image analysis since it allows to investigate image features independent of illumination. However, analytical wavelets are only one dimensional and therefore only adaptable for image analysis by a tensor product approach.

Steerable filters are directional filters. Directional filters are used in image processing for example in texture analysis, edge detection, image data compression, motion analysis and image enhancement. (See [18] for references.) For these tasks filters with arbitrary directions are often required. A steerable filter in an arbitrary direction is synthesized as a linear combination of a set of basis filters [18] - in contrast to other directional filters like Gabor or Morlet wavelets, where a separate filter is necessary for each orientation. Of course this restricts the number of possible directions for non steerable directional filters. However, steerable filters are usually derived as steerable pyramids that yield only discrete filters.

Our aim is to transfer the concept of analytic wavelets into arbitrary higher dimensions using the Riesz transform which generates steerable filters.

Our approach is new in that we use a wavelet frame construction that works in n-D for any $n \in \mathbb{N}$. Furthermore this approach is intrinsically continuous while still allowing for an efficient discrete implementation. The benefit of our approach is that we retrieve phase information which is localized in space and scale from a signal of arbitrary dimension using filters which are steerable multiwavelets. The phase we derive in this multi-scale approach is better interpretable than the one derived by the monogenic signal, because it gives phase information on all scales separately, whereas, in contrast, the phase derived from the monogenic signal is an average over all scales.

The fact that we use wavelets as filters enables us to use the advantageous properties of continuous wavelets to analyze the properties of our filters - continuity, decay, vanishing moments and localization carry information about important properties as for example the approximation order.

We noted that the Riesz transforms of our wavelets are steerable filters. Indeed they can be viewed as steerable pyramids. But this connection is not one way – using higher Riesz transforms we state a method to find the operators which generate steerable filters. In this way steerable filters are freed of the discrete setting of steerable pyramids and can be viewed as steerable operators which may act on radial wavelet frames to generate steerable wavelets as well as on any radial filters that generate a frame.

To show that the Riesz transform of a wavelet frame yields a multiwavelet frame and to investigate the space that monogenic wavelets generate we consider Clifford frames on non-commutative Clifford-Hilbert modules. Since frames and multiwavelet frames can be viewed as Clifford frames this setting allows us to proof that the Riesz transform preserves Clifford frames. The same means enable us to study the space generated by monogenic wavelet frames.

More interesting than a historical account of this field is an account of the development in this exciting area of research – driven by several groups of strong researchers – that occured during the time in which this thesis was written. (An account of the earlier development of this field is given in our paper [24]. Especially the works of Elias Stein and Guido Weiss [45, 46] and the work of Michael Felsberg[15] should be noted in this context.)

In 2009 Michael Unser and Dimitri Van De Ville published a paper [52] in which they demonstrated an implementation of the Riesz transform in 2-D using an orthogonal basis of approximately radial wavelets. For the proof that the Riesz transform of an wavelet frame is once again a wavelet frame this paper refers to our paper [23] which was at that time in review process.

In [38] Juan Romero et. al. gave a method to construct radial tight wavelet frames and gave examples of radial wavelet frames with $|\det A|$ radial mother wavelets, where A is the dilation matrix.

A first part of this thesis was published in [23] and [24] which contains the implementation of the Riesz transforms in arbitrary dimension using tight wavelet frames with a single radial mother wavelet and the theory of Clifford frames to proof the frame property of the Riesz transforms.

These papers contain several new contributions:
In contrast to Michael Unser, we use wavelet frames in arbitrary dimensions. These frames are perfectly radial, which improves the quality of the derived phase and allows for proper steerability.
In contrast to Romero et. al. we use a tight wavelet frame with exactly one mother wavelet. This is less redundant and allows for an easier application of the Riesz transform. Furthermore, we proof that the Riesz transform maps wavelet frames to multiwavelet frames. It is this property that makes it possible to design perfect reconstruction filter banks. In several other publications in this field e.g. [42, 51, 53, 4, 48, 49] our article [24] was cited.

In [24] we proposed to study higher Riesz transforms to generate steerable wavelets.

A first step in this direction were the papers [51] and [53] where Michael Unser et. al. introduced higher order Riesz transforms, which correspond to a composition of higher Riesz transforms we propose in this thesis.

The chapter on higher Riesz transforms in this thesis is concerned with giving a comprehensive theoretical background on higher Riesz transforms and steerable filters. Using this background we can discern the minimal higher Riesz transforms which are steerable and give the means to combine these adaptively to higher Riesz transforms with more interesting geometrical features. This is important to gain flexibility while limiting redundancy. The number of partial higher Riesz transforms linearly increases the runtime and memory use of the algorithms.

Another advantage of our approach is that we can use our theoretic knowledge – namely that the steerability is a manifestation of a representation of the rotation group – to gain insights into the higher order Riesz transforms of Michael Unser's. We show that these higher order Riesz transforms come in two sets – the filters of odd and even order – in which all sets of higher Riesz transforms of lower order are contained in those of higher order. As a consequence the redundancy of these higher order Riesz transforms rapidly increases as the order increases.

Furthermore we were able to show a connection between higher Riesz transforms of wavelets and the steerable pyramids described in [18]. It is easily seen that the implementation of the higher Riesz transform of a wavelet frame is a steerable pyramid. On the other hand we could show that steerable pyramids correspond to higher Riesz transforms applied to filters which form a perfect reconstruction filter bank.

1.2 Contents of the chapters

Chapter 1 provides a short summary of basic facts on harmonic analysis and analytical signals which will prove useful in the remainder of this thesis.

Chapter 2 introduces the Riesz transform. The first part of this chapter section 2.1 introduces the Riesz transform as a unique extension of the Hilbert transform, states the most important classical results on the Riesz transform, gives an interpretation of the term "analytical" in the name "analytical signal" and states that Riesz transforms are steerable.

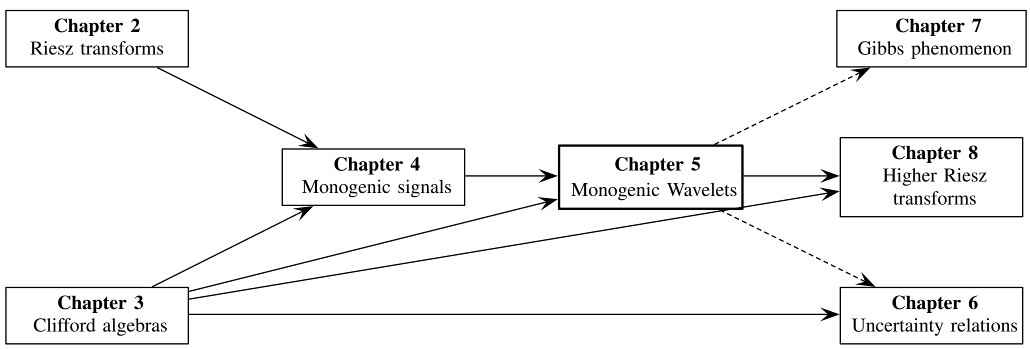


Figure 1.1: Interrelation of the chapters

The second part of this chapter, section 2.2, presents new results. We define the Riesz transform for two classes of distributions – distributions modulo polynomials (Definition 2.2.10) and tempered distributions whose support in the Fourier domain does not contain the origin (Theorem 2.2.13). As a application we state a novel extension of the Bedrosian identity to Hilbert transforms of distributions (Theorem 2.2.18). Our definition of the Riesz transform of distributions is especially well suited for this task — indeed the prerequisites for the Bedrosian identity are all that is required for the Hilbert transform of the involved distributions to be well defined. By means of a certain new operator that maps one dimensional tempered distributions to n dimensional tempered distributions (Proposition 2.2.14) we state a new connection between the Hilbert and the Riesz transform (Theorem 2.2.15). From this connection we then obtain a Bedrosian identity for Riesz transforms of certain distributions (Theorem 2.2.19).

Chapter 3 introduces Clifford algebras and Clifford Hilbert-modules — the language in which many of the following results are formulated.

Section 3.1 gives a short introduction to Clifford algebra.

Section 3.2 introduces Clifford modules to provide the setting in which we then introduce some already known facts about functional analysis in Clifford-Hilbert modules.

Section 3.2.3 considers for the first time systematically left-linear operators on Clifford-Hilbert modules. This consideration results in new insights that are stated in Theorem 3.2.7 – Corollary 3.2.13. These insights are essential for the definition of the monogenic signal in chapter 4, for deriving uncertainty relations in chapter 6 and for the development of Clifford frames in section 5.2.

The final part of the chapter treats some basics of Clifford analysis which extends classical complex analysis to higher dimensions. The monogenic signal is closely connected to Clifford analysis as revealed in chapter 4.

In chapter 4 we formulate the monogenic signal as an extension of the analytical signal to higher dimensions and examine the properties of this monogenic signal.

Section 4.1 establishes the hypercomplex Riesz transform on which the monogenic signal is based. The first step is the definition of the hypercomplex Riesz transform R in Definition 4.1.1 that extends the Hilbert transform $i\mathscr{H}$. What is novel about this definition is that we define the Riesz transform as a left linear-operator on a Clifford-Hilbert module. As a consequence the (hypercomplex) Riesz transform we define is unitary and self-adjoint as is shown in Theorem 4.1.4. It follows that the hypercomplex Riesz transform is its own inverse.

Section 4.2 shows that the hypercomplex Riesz transform is closely connected to differential operators. This novel result is stated in Theorem 4.2.1.

We define the monogenic signal in section 4.3.

In section 4.4 we state the decomposition of the monogenic signal into phase, amplitude, and phase direction. The later provides the preferred direction of a signal. From the phase we derive an instantaneous frequency for the monogenic signal.

By the properties of the Riesz transforms defined in Definition 4.1.1 it follows that the monogenic signal is generated by a projection $\mathscr{M} : f \mapsto f_m$ (Theorem 4.3.3) that is closely connected to Clifford analysis. We state these connections in section 4.5.

The final section 4.6 shows that the monogenic signal is uniquely defined as an extension of the analytical signal by the condition that it is compatible with rotations – which generate the symmetry group we have to deal with in image analysis.

Some of the results of this chapter have been published in the peer reviewed article [24] and in [23].

The agenda of chapter 5 is to implement the monogenic signal – given by a hypercomplex singular integral operator - via multi-wavelet frames which amounts to a multiplication with a set of explicitly known filters in the Fourier domain.

This is the central chapter of this thesis - all previous chapters set the background for what we achieve in this chapter and all further chapters extend what we do in this chapter.

The first section 5.1 explains why wavelet frames are the tool of choice for the implementation of the monogenic signal. Example 5.1.2 illustrates our approach using the example of analytical wavelet frames and sets the roadmap for sections 2 and 3.

Section 5.2 extends the concept of frames on Hilbert spaces to the setting of Clifford-Hilbert modules – the space in which the monogenic signal lives – introduced in section 3.2. The concept of Clifford frames – to use Clifford algebra valued frame coefficients to achieve a frame decomposition – appears to be entirely new.

Section 5.3 defines the notion of a hypercomplex wavelet transform which is compatible with Clifford frames (Definition 5.3.1) and the monogenic signal. Theorem 5.3.3 shows that we can derive a monogenic wavelet frame from a standard wavelet frame for $L^2(\mathbb{R}^n)$ using a surjective hypercomplex operator.

In section 5.4 we state that under mild conditions the decay rate of the wavelet frames is preserved by the Riesz transform.

Section 5.5 is dedicated to the search for suitable wavelet frames. First, we introduce results which reduce the problem of constructing wavelet frames to the problem of constructing Riesz partitions of unity. Theorem 5.5.7 provides a new way to find such Riesz partitions of unity. An explicit construction is then given in Example 5.5.1. Theorem 5.5.8 shows how to derive Riesz partitions of unity from agiven compactly supported wavelet orthonormal bases of $L^2(\mathbb{R})$.

Finally, section 5.6 presents the implementation of the wavelet frames of Example 5.5.1 as an imageJ plugin. The software and the applications have been the topic of the diploma thesis of Martin Storath [47] under supervision of the author.

The result is a tight steerable multiwavelet frame in arbitrary dimensions that yields a decomposition into amplitude, phase and a direction which is local with respect to location and scale.

The tight multi-wavelet frames do have a simple and explicit form in the Fourier domain, whence all algorithms are implemented in the frequency domain. The compact support of the wavelet frames in the Fourier domain allows a fast filter-bank implementation with perfect reconstruction property. The Shannon sampling theorem ensures loss-less down-sampling after each filter step.

Some of the results of this chapter have been published in the peer reviewed article [24] and in [23].

In chapter 6 we examine uncertainty relations for the Riesz transforms. Uncertainty relations quantify the joint localisation in time or space domain and Fourier domain. For that reason they are an important tool in the design of wavelets. In chapter 5 we implement the Hilbert and the Riesz transform via wavelets. Therefore it will be the topic of chapter 6 to compute the effect of the Hilbert and the Riesz transform on the Weyl-Heisenberg uncertainty relation.

In Theorem 6.2.3 we prove the new result that under mild assumptions the Weyl-Heisenberg uncertainty relation is invariant under the Hilbert transform.

Since the Riesz transform R is unitary whereas the partial Riesz transforms R_α are not we consider a single uncertainty relation for the Riesz transform rather than a set of uncertainty relations for the partial Riesz transforms. This requires a new kind of uncertainty relation for vector valued functions.

In section 6.3 we pursue this goal along two different paths: a classical approach Theorem 6.3.2 based on the Cauchy-Schwartz inequality and an uncertainty relation on Clifford-Hilbert modules Theorem 6.3.3. Both paths yield novel uncertainty relations.

In contrast to the set of classical Weyl-Heisenberg uncertainty relations in higher dimensions stated in Corollary 6.3.1 the Weyl-Heisenberg uncertainty relation we derive in Theorem 6.3.10 is invariant under rotations. This property of invariance with respect to rotation makes them easily interpretable in the context of image processing.

In Theorem 6.3.12 we use our results to derive a novel Weyl-Heisenberg uncertainty relation for the Riesz transform and to compute the effect of the Riesz transform on the Weyl-Heisenberg uncertainty relation Theorem 6.3.10.

The Chapter will end with a new affine uncertainty relation Theorem 6.3.13 as a further example for an application of Theorem 6.3.3.

In chapter 7 we state sufficient conditions for the Gibbs phenomenon to occur in a wavelet frame decomposition. The Gibbs phenomenon describes the appearance of over- and undershoots of a wavelet approximation near a jump discontinuity. We examine the existence of the Gibbs phenomenon for wavelet frames which are based on a certain kind of multiresolution analysis, i.e., for which a suitable scaling function exists.

Chapter 8 is dedicated to the study of higher Riesz transforms. The operator resulting from successively applying several partial Riesz transforms is called a higher Riesz transform.

The steerability of the Riesz transform corresponds to a unitary representation of the rotation group on the set of Riesz transformed functions. In section 8.2 we show that for $n > 1$ there are an infinite number of such unitary representations of the rotation group based on the spherical harmonics. From these we derive steerable unitary left linear operators on Clifford-Hilbert modules - the higher Riesz transforms. These higher Riesz transforms are the smallest possible sets of partial higher Riesz transforms which are invertible and steerable.

Combining these minimal higher Riesz transforms in Corollary 8.2.18 we derive higher Riesz transforms which are geometrically richer at the cost of higher redundancy.

We use the theory of Clifford frames and hypercomplex wavelets we developed for the Riesz transform in chapter 5 to implement the higher Riesz transforms. The higher Riesz transforms give rise to a higher monogenic signal, which yields a decomposition into amplitude, phase, and geometrical information. Theorem 8.2.15 states the connection of the higher monogenic signal to certain generalized Cauchy Riemann equations derived from higher Dirac operators of higher order which are the square-root of a power of the Laplace operator.

In Remark 8.2.19 we learn how to construct higher Riesz transforms which contain the directional information of higher derivatives. This approach is similar to the steerable pyramids of Freeman and Adelson in [18] – indeed in Remark 8.2.22 we show that these steerable pyramids have a close connection to higher Riesz transforms.

In section 8.3 we give explicit constructions of higher Riesz transforms for the interesting cases $n = 2$ and $n = 3$.

The Appendix A states some conventions and give some further details on distributions.

1.3 Preliminaries

In this section we will state some basic facts of harmonic analysis that we often will use later on.

1.3.1 Steerable filters

Steerable filters were introduced by Freeman and Adelson in [18]. A steerable filter is a function which can be steered, i.e. arbitrary rotated versions of the function can be computed by a linear combination of a certain finite set of rotated versions of the function.

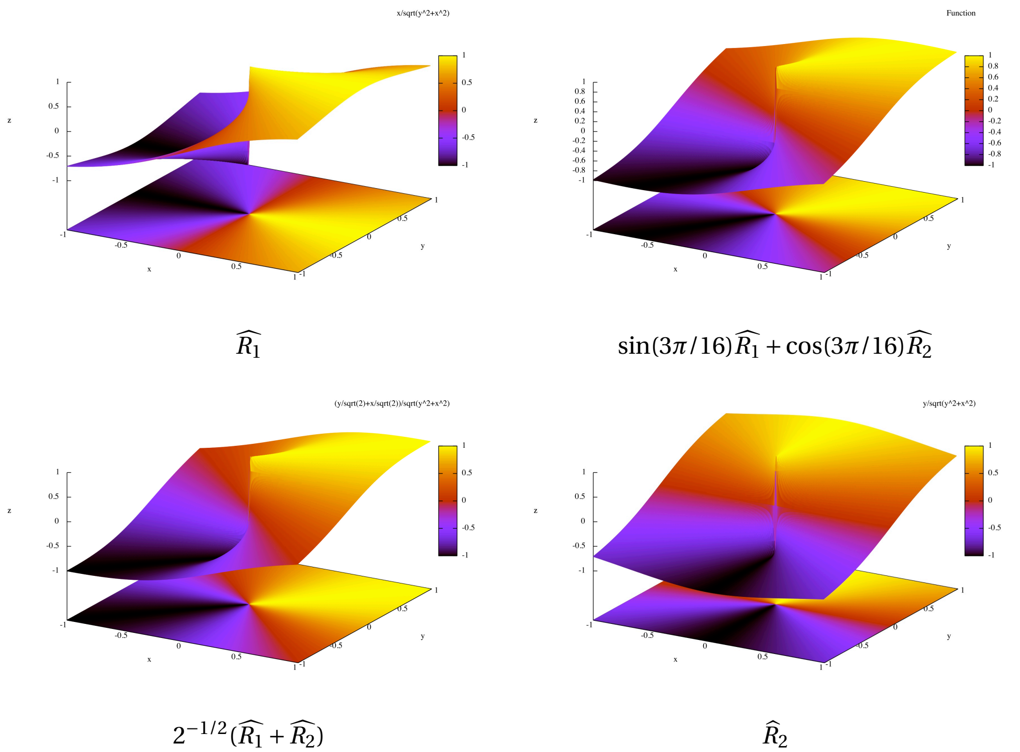


Figure 1.2: Example of steerable filters: the basis filters are R_1, R_2 the linear combinations are rotated versions of the filters.

Definition 1.3.1 (Steerable filter)
LET $\{f_a\}_{a\in S^{n-1}}$ *be a set of directed filters.*

THEN f_a *is called a* **steerable filter**, **iff** *the steerable filter with respect to an arbitrary direction* $a \in S^{n-1}$ *can be written as a linear combination of filters with respect to a set of basis directions* $\{a_l\}_{l=1,\ldots,k}$,. *That is for every* $a \in S^{n-1}$ *there exists a vector* $v_a = (v_{a,1}, \ldots, v_{a,l}) \in \mathbb{R}^k$ *such that*

$$f_a = \sum_{l=1}^{k} v_{a,l} f_{a_l}.$$

1.3.2 Tempered distributions

We will start with some basic facts about tempered distributions. A more thoroughly introduction is given in the appendix, section A.2.

Definition 1.3.2 (Schwartz Class)
LET $f \in C^\infty(\mathbb{R}^n)$, $n \in \mathbb{N}$. *Furthermore let* $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}_0^n$ *be a multi-index. As usual we write*

$$\partial^\alpha f(x) = \frac{\partial^{|\alpha|}}{\partial^{\alpha_1} x_1 \cdots \partial^{\alpha_n} x_n} f(x),$$

where $|\alpha| = \sum_{l=1}^n \alpha_l$. *Let*

$$\rho_{j,k}(f) = \sup_{x \in \mathbb{R}^n,\, |\alpha| \le j} (1 + |x|)^k |\partial^\alpha f(x)|.$$

THEN *the* **Schwartz class** *of rapidly decreasing functions* $\mathscr{S}(\mathbb{R}^n)$ *is defined as*

$$\mathscr{S}(\mathbb{R}^n) := \{f \in C^\infty(\mathbb{R}^n) : \forall j, k \in \mathbb{N}_0 \exists C_{j,k} \in \mathbb{R}^+ : \rho_{j,k}(f) = C_{j,k} < \infty\}.$$

Theorem 1.3.3
$\mathscr{S}(\mathbb{R}^n)$ *is a Fréchet space with the set of seminorms* $\rho_{j,k}$.
$\mathscr{S}(\mathbb{R}^n)$ *is a dense subspace of* $L^p(\mathbb{R}^n)$, $1 \le p < \infty$.

Proof. For a proof see [37] V.3. □

Definition 1.3.4 (Tempered distributions)
Elements of the dual space $\mathscr{S}'(\mathbb{R}^n)$ *of* $\mathscr{S}(\mathbb{R}^n)$ *are called* **tempered distributions**.
The support of $f \in \mathscr{S}'(\mathbb{R}^n)$ *is defined by*

$$\operatorname{supp}(f) := \bigcap \{K \subset \mathbb{R}^n \text{ compact} : \phi \in \mathscr{S}(\mathbb{R}^n),\ \operatorname{supp}(\phi) \subseteq K^c \Rightarrow f(\phi) = 0\}.$$

Definition 1.3.5 (Derivatives of distributions)
LET $f \in \mathscr{S}'(\mathbb{R}^n)$ *and let* $\boldsymbol{\alpha} \in \mathbb{N}_0^n$.
THEN *the derivative* $\partial^{\boldsymbol{\alpha}} f \in \mathscr{S}'(\mathbb{R}^n)$ *is defined as*

$$\partial^{\boldsymbol{\alpha}} f(\phi) := (-1)^{|\boldsymbol{\alpha}|} f(\partial^{\boldsymbol{\alpha}} \phi),\ \forall \phi \in \mathscr{S}(\mathbb{R}^n).$$

Corollary 1.3.6
The derivative of a distribution

$$\partial^{\boldsymbol{\alpha}} : \mathscr{S}'(\mathbb{R}^n) \to \mathscr{S}'(\mathbb{R}^n)$$

as given in Definition 1.3.5 is well defined.

Proof. For a proof see for example [41] 6.12. □

Definition 1.3.7 (The direct product of distributions)
LET $n, m \in \mathbb{N}$, $f \in \mathscr{S}'(\mathbb{R}^n)$, *and* $g \in \mathscr{S}'(\mathbb{R}^m)$.
THEN *the* **direct product** $f \times g \in \mathscr{S}'(\mathbb{R}^{n+m})$ *of* f *and* g *is defined as*

$$(f \times g)(\phi) := f(g(\phi)) = g(f(\phi)),\ \forall \phi \in \mathscr{S}(\mathbb{R}^n \times \mathbb{R}^m).$$

Definition 1.3.8 (Convolution of distributions)
LET $f, g \in \mathscr{S}'(\mathbb{R}^n)$.

THEN *the* **convolution** $f * g \in \mathscr{S}'(\mathbb{R}^n)$ *of* f *and* g *is defined by*

$$f * g(\phi) = f \times g\big(\phi(\,\cdot\,_f + \,\cdot\,_g)\big),\ \forall \phi \in \mathscr{S}(\mathbb{R}^n).$$

Remark 1.3.9 (Well-definedness of direct product and convolution)
The direct product of distributions given in Definition 1.3.7 is well-defined and commutative. (See Theorem A.2.7.) The convolution in Definition 1.3.8 is well defined if at least one of the tempered distributions has compact support. A proof of this fact can be found in the appendix Theorem A.2.9.

Theorem 1.3.10 (Product of distributions)
LET $g \in \mathscr{S}'(\mathbb{R}^n)$ *and let* $f \in C^\infty(\mathbb{R}^n)$ *such that*

$$\forall \alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}_0^n\, \exists k_\alpha \in \mathbb{N}_0, C_\alpha > 0 : |(\partial^\alpha f)(x)| \le C_\alpha (1 + |x|)^{k_\alpha},\ \forall x \in \mathbb{R}^n.$$

THEN *the* **product** $fg \in \mathscr{S}'(\mathbb{R}^n)$ *is well defined by*

$$fg(\phi) = g(f\phi),\ \forall \phi \in \mathscr{S}(\mathbb{R}^n).$$

1.3.3 The Fourier transform

Definition 1.3.11 (Fourier transform)
We define the **Fourier transform** *of a function* $f \in \mathscr{S}(\mathbb{R}^n)$ *by*

$$\mathscr{F}(f)(\xi) = \widehat{f}(\xi) := \int_{\mathbb{R}^n} f(x) e^{-2\pi i x\cdot\xi}\, dx = \int_{\mathbb{R}^n} f(x) e^{-2\pi i \langle x, \xi\rangle_{\mathbb{R}^n}}\, dx, \text{ f.a.a. } \xi \in \mathbb{R}^n.$$

Theorem 1.3.12
The Fourier transform is an isomorphism on $\mathscr{S}(\mathbb{R}^n)$. *Its inverse is given by*

$$\mathscr{F}^{-1}(f)(x) := \int_{\mathbb{R}^n} f(\xi) e^{2\pi i x\cdot\xi}\, d\xi, \text{ f.a.a. } x \in \mathbb{R}^n.$$

The Fourier transform of a function in $L^p(\mathbb{R}^n)$, $1 \le p < \infty$, *is defined by extension of this definition to* $L^p(\mathbb{R}^n)$ *and denoted by* $\mathscr{F}$. *By the Plancherel theorem the Fourier transform is an isometry on* $L^2(\mathbb{R}^n)$.

The Fourier transform $\mathscr{F} : \mathscr{S}'(\mathbb{R}^n) \to \mathscr{S}'(\mathbb{R}^n)$ *of a tempered distribution* $f \in \mathscr{S}'(\mathbb{R}^n)$ *is defined as* $\widehat{f}(\phi) = f(\widehat{\phi})$, $\forall \phi \in \mathscr{S}(\mathbb{R}^n)$. *It is an isomorphism on* $\mathscr{S}'(\mathbb{R}^n)$.

Proof. For the proof see [40], Chapter 7, and Theorem 7.4-1in [58] . □

Theorem 1.3.13 (Plancherel)
LET $f \in L^2(\mathbb{R}^n)$.

THEN

$$\|f\|_{L^2(\mathbb{R}^n)} = \|\widehat{f}\|_{L^2(\mathbb{R}^n)}.$$

Proof. A proof can be found in [40]. □

Corollary 1.3.14 (Parsevals formula)
LET $f, g \in L^2(\mathbb{R}^n)$.
THEN

$$\langle f, g\rangle = \langle \widehat{f}, \widehat{g}\rangle.$$

Proof. A proof can be found in [6]. □

Theorem 1.3.15 (Fourier transform and convolution)
LET $f, g \in \mathscr{S}'(\mathbb{R}^n)$ *such that* $\widehat{f}$ *has compact support.*
THEN

$$\widehat{fg} = \widehat{f} * \widehat{g}.$$

Proof. The proof is given in the appendix Theorem A.2.14. □

Theorem 1.3.16 (Poisson's Summation Formula)

(i) LET $f : \mathbb{R}^n \to \mathbb{R}$ *such that*

$$\exists \epsilon > 0,\ C > 0 : |f(x)| \leq C(1+|x|)^{-n-\epsilon} \wedge |\widehat{f}(\xi)| \leq C(1+|\xi|)^{-n-\epsilon}.$$

THEN

$$\sum_{k\in\mathbb{Z}^n} f(x+k) = \sum_{k\in\mathbb{Z}^n} \widehat{f}(k)e^{2\pi i k\cdot x},$$

holds pointwise for all $x \in \mathbb{R}^n$*, and both sums converge absolutely for all* $x \in \mathbb{R}^n$.

(ii) LET $f \in L^2(\mathbb{R}^n)$ *such that* $\sum_{k\in\mathbb{Z}^n} f(x+k) \in L^2([0,1]^n)$ *and* $\sum_{k\in\mathbb{Z}^n} |\widehat{f}(n)|^2 < \infty$. THEN

$$\sum_{k\in\mathbb{Z}^n} f(x+k) = \sum_{k\in\mathbb{Z}^n} \widehat{f}(k)e^{2\pi i k\cdot x},$$

holds almost everywhere, and both sums converge in $L^2(\mathbb{R}^n)$.

Proof. For a proof see [20] Chapter 1.4. □

Theorem 1.3.17 (Fourier transforms and derivatives)
LET $f \in \mathscr{S}(\mathbb{R}^n)$ *and let* $\boldsymbol{\alpha} \in \mathbb{N}_0^n$.
THEN

$$\mathscr{F}(\partial^{\boldsymbol{\alpha}} f)(\xi) = (2\pi i\xi)^{\boldsymbol{\alpha}} \widehat{f}(\xi),\ \forall \xi \in \mathbb{R}^n,$$

and

$$\mathscr{F}\big((-2\pi i\xi)^{\boldsymbol{\alpha}} f\big)(\xi) = \partial^{\boldsymbol{\alpha}} \widehat{f}(\xi).$$

Proof. For a proof see for example [20]. □

Definition 1.3.18 (The Laplace operator)
LET $f \in \mathscr{S}(\mathbb{R}^n)$. *The* **Laplace operator** *is given by*

$$\triangle f(x) := \sum_{\alpha=1}^{n} \frac{\partial^2}{\partial x_\alpha^2} f(x).$$

By Definition 1.3.5, it can be extended to $f \in \mathscr{S}'(\mathbb{R}^n)$.

Definition 1.3.19 (Sobolev and Bessel potential spaces)
LET $k \in \mathbb{N}$, $p \in [1,\infty[$. *The* **Sobolev spaces** *are defined as*

$$W^{k,p}(\mathbb{R}^n) = \{f \in \mathscr{S}'(\mathbb{R}^n) : \partial^{\boldsymbol{\alpha}} f \in L^p(\mathbb{R}^n),\ \forall \boldsymbol{\alpha} \in \mathbb{N}_0^n,\ |\boldsymbol{\alpha}| \leq k\}.$$

Let $s \in \mathbb{R}$. *The* **Bessel potential** *spaces are given by*

$$W^s(\mathbb{R}^n) := \{f \in \mathscr{S}'(\mathbb{R}^n) : \int_{R^n} \widehat{f}^2(\xi)(1+|\xi|^2)^s < \infty\}.$$

Theorem 1.3.20
LET $f \in \mathscr{S}'(\mathbb{R}^n)$, $k \in \mathbb{N}$.
THEN $\partial^{\boldsymbol{\alpha}} f \in L^2(\mathbb{R}^n),\ \forall |\boldsymbol{\alpha}| \leq k \Leftrightarrow \int_{\mathbb{R}^n} |\widehat{f}(\xi)|^2(1+|\xi|^2)^k < \infty$.
As a consequence

$$W^{k,2}(\mathbb{R}^n) = W^k(\mathbb{R}^n).$$

Proof. For a proof see [20], Lemma 1.2.3. □

1.3.4 Basic operators of harmonic analysis

We will now define some basic operators of harmonic analysis.

Definition 1.3.21 (Translation, modulation, dilation and rotation)
LET $t, b \in \mathbb{R}^n$, $d, d_i \in \mathbb{R}$, $\forall 1 \leq i \leq n, i \in \mathbb{N}$ *and* $f : \mathbb{R}^n \mapsto \mathbb{C}$.
THEN *we define the* **translation** *operator by*

$$T_t f(x) := f(x-t),$$

and the **modulation** *operator by*

$$M_b f(x) := e^{2\pi i b \cdot x} f(x).$$

We define the operation of a **rotation** $\rho \in SO(n)$ *on a function* f *by*

$$\rho f(x) = f(\rho x).$$

We define a **dilation** *by a matrix* A *with eigenvalues of modulus greater than one via*

$$Af(x) := |\det(A)|^{1/2} f(Ax).$$

This definition is usually too general so that we have to restrict to certain types of dilation.
The **isotropic dilation** *is defined by*

$$D_d f(x) := |d|^{n/2} f(dx).$$

In wavelet theory this is often combined with a rotation ρ *which gives what we will call* **rotated dilation**

$$\mathfrak{D}_d f(x) := D_d \rho f(x) = |d|^{n/2} f(d\rho x).$$

In product space settings where rotations are not considered, a **non-isotropic dilation** *is useful which will be defined as*

$$\mathscr{D}_{\{d_i\}_{i=1}^n} f(x) := \left(\prod_{i=1}^n d_i\right)^{1/2} f(\operatorname{diag}(d_1,\dots,d_n)x),$$

where $\operatorname{diag}(d_1,\dots,d_n)$ *is the diagonal* $n\times n$*-matrix with entries* $d_1,\dots,d_n \in \mathbb{R}$.

Theorem 1.3.22 (Fourier transforms of translation, modulation, dilation and rotation)
LET $t, b \in \mathbb{R}^n$, $d, d_i \in \mathbb{R}$, $\forall 1 \le i \le n, i \in \mathbb{N}$ *and* $f : \mathbb{R}^n \mapsto \mathbb{C}$. *Let* $f \in L^p(\mathbb{R}^n)$ *or* $f \in \mathscr{S}(\mathbb{R}^n)$.
THEN *modulation and translation are dual with respect to the Fourier transform in the sense that*

$$\widehat{T_t f} = M_{-t}\widehat{f}$$

and

$$\widehat{M_b f} = T_b \widehat{f}.$$

Both are unitary operators with respect to the L^2*-norm, which is clear for* M_b *and hence by the Plancherel theorem for* T_t.
The inverse of the rotation operator is the operator corresponding to the transposed matrix which is the inverse of the rotation matrix and

$$\mathscr{F}(\rho f) = \rho\mathscr{F}(f),\ \forall f \in L^2(\mathbb{R}^n).$$

The Fourier transform of the isotropic dilation is

$$\widehat{D_d f} = D_{d^{-1}}\widehat{f} = D_d^{-1}\widehat{f},\ \forall f \in L^p(\mathbb{R}^n).$$

The Fourier transform of the rotated dilation is

$$\widehat{\mathfrak{D}_d f} = \widehat{D_d \rho f} = D_{d^{-1}}\rho\widehat{f} = (\mathfrak{D}^T)^{-1}\widehat{f},\ \forall f \in L^p(\mathbb{R}^n).$$

Proof. Let $f \in L^p(\mathbb{R}^n)$ or $f \in \mathscr{S}(\mathbb{R}^n)$. Modulation and translation are dual since

$$\begin{aligned}\mathscr{F}(T_t f)(\xi) &= \int_{\mathbb{R}^n} f(x-t)e^{-2\pi i\langle x,\xi\rangle}dx = \int_{\mathbb{R}^n} f(x)e^{-2\pi i\langle x+t,\xi\rangle}dx = e^{2\pi i\langle -t,\xi\rangle}\int_{\mathbb{R}^n} f(x)e^{-2\pi i\langle x,\xi\rangle}dx\\ &= M_{-t}\mathscr{F}(f)(\xi)\end{aligned}$$

and

$$\begin{aligned}\mathscr{F}(M_b f)(\xi) &= \int_{\mathbb{R}^n} f(x)e^{2\pi i\langle b,x\rangle}e^{-2\pi i\langle x,\xi\rangle}dx = \int_{\mathbb{R}^n} f(x)e^{-2\pi i\langle x,\xi-b\rangle}dx\\ &= T_b\mathscr{F}(f)(\xi),\ \forall \xi \in \mathbb{R}^n.\end{aligned}$$

All the other operators are of the form $Af(\,\cdot\,) := f(A\,\cdot\,)$, where $A \in \mathrm{Gl}(n)$ is an invertible matrix. In general it holds that

$$\begin{aligned}\mathscr{F}(Af)(\xi) &= \int_{\mathbb{R}^n} f(Ax)e^{-2\pi i\langle x,\xi\rangle}dx = \int_{\mathbb{R}^n} f(x)e^{-2\pi i\langle A^{-1}x,\xi\rangle}|\det(A^{-1})|dx\\ &= \int_{\mathbb{R}^n} f(x)e^{-2\pi i\langle x,(A^{-1})^*\xi\rangle}|\det(A^{-1})|dx\\ &= |\det(A^{-1})|\big(A^{-1}\big)^*\mathscr{F}(f)(\xi).\end{aligned}$$

□

1.3.5 Wavelets and frames

Definition 1.3.23 (Wavelet transform)
LET $\psi \in L^2(\mathbb{R}^n)$ *and let* D *be a dilation matrix.*

THEN *the set of dilated and translated versions of* ψ

$$\{D^j T_k \psi\}_{j,k}$$

is called a **wavelet system**, *where* $j \in \mathbb{Z}$ *and* $k \in \mathbb{Z}^n$ *or* $j \in \mathbb{R}$ *and* $k \in \mathbb{R}^n$ *with* **mother wavelet** ψ.

LET $f \in L^2(\mathbb{R}^n)$.

THEN

$$W_\psi(f) := \left\{\langle f, D^j T_k \psi\rangle\right\}_{j,k}$$

is called the **wavelet transform** *of* f.

Definition 1.3.24 (Frame)
LET H *be a Hilbert space, let* I *be a countable set and* $\{\psi_k\}_{k\in I} \subset H$.

THEN $\{\psi_k\}_{k\in I}$ *is called a* **frame** *for* H *iff* $\exists 0 < A \le B < \infty$ *such that* $\forall f \in H$ *the* **frame inequality**

$$A\|f\|^2 \le \sum_{k\in I} |\langle f, \psi_k\rangle|^2 \le B\|f\|^2$$

holds. In this case A *and* B *are called the* **lower** *respectively the* **upper frame bound**. $\{\psi_k\}_{k\in I}$ *is called a* **tight** *frame iff* $A = B$.

Definition 1.3.25 (Wavelet frame)
A wavelet system is called a **wavelet frame** *if it is a frame.*

Theorem 1.3.26 (Dual frame and frame decomposition)
LET $\{\psi_k\}_{k\in I} \subset H$ *be a frame for a Hilbert space* H.

THEN *there exists a frame* $\{\phi_k\}_{k\in I} \subset H$ *called a* **dual** *frame such that* $\forall f \in H$ *the* **frame decomposition**

$$f = \sum_{k\in I} \langle f, \psi_k\rangle \phi_k = \sum_{k\in I} \langle f, \phi_k\rangle \psi_k$$

holds.

Proof. For a proof and for further details on frame theory see [10]. □

The image of a frame under a surjective operator yields a frame:

Theorem 1.3.27 (Frames and operators)
LET H, G *be Hilbert spaces. Let* $\{f_k\}_{k\in\mathbb{N}}$ *be a frame for* H *with bounds* $0 < A \le B < \infty$ *and let* $U : H \to G$ *be a bounded surjective operator.*

THEN $\{Uf_k\}_{k\in\mathbb{N}}$ *is a frame for* G *with frame bounds* $A\|U^\dagger\|^{-2}, B\|U\|^2$. *(Here* † *denotes the pseudo-inverse.)*

Proof. This is Theorem 5.3.2 in [10]. □

1.3.6 Representations of groups and algebras

Definition 1.3.28 (Representation)
LET *G be a group or an algebra and let X be a linear space. A* **representation** *of G in X is a homomorphism $\mathscr{R}$ from G to the vector space $L(X)$ of linear, bijective mappings from X to X. Let $g, h \in G$ and let $e \in G$ be the identity.*

THEN

$$\mathscr{R}(g)\mathscr{R}(h) = \mathscr{R}(gh),$$
$$\mathscr{R}(e) = \mathrm{Id}_{L(X)},$$

and in the case that G is an algebra, we have in addition, that

$$\mathscr{R}(g+h) = \mathscr{R}(g) + \mathscr{R}(h).$$

A representation is called **reducible** *if there exists a subspace V of X that is invariant under the representation $\mathscr{R}$, i.e.,*

$$\mathscr{R}(g)\, v \in V,\ \forall\, v \in V,\ g \in G.$$

Then the restriction of $\mathscr{R}$ to V is called a **subrepresentation** *of $\mathscr{R}$.*

The **factor representation** *is the representation on the factor space X/V naturally given by $\mathscr{R}$.*

A representation is called **unitary** *if the space X is a Hilbert space and $\mathscr{R}(g)$ is unitary for all $g \in G$.*

Remark 1.3.29
LET *$\mathscr{R}$ be a unitary reducible representation on a vector space X with invariant subspace V.*

THEN *$\mathscr{R}|_{V^\perp}$ is a unitary representation.*

1.4 Analytical signals

1.4.1 The Hilbert transform

We define analytical signals using the Hilbert transform, so we will first have a look at the Hilbert transform itself.

Definition 1.4.1 (The Hilbert transform)
LET $f \in L^p(\mathbb{R})$, *where* $1 < p < \infty$.
THEN *the* **Hilbert transform**

$$\mathscr{H} : L^p(\mathbb{R}^n) \mapsto L^p(\mathbb{R}^n),$$

is defined as the following Cauchy principal value:

$$\mathscr{H} f(x) := \lim_{\delta \to 0} \frac{1}{\pi} \int_{|y|>\delta} \frac{f(x-y)}{y} dy.$$

Theorem 1.4.2 (Properties of the Hilbert transform)
LET $f \in L^p(\mathbb{R})$, *where* $1 < p < \infty$.
THEN *the Hilbert transform*

(i) anti-commutes with reflection, i.e. $\mathscr{H}(f(-\,\cdot\,))(x) = -\mathscr{H}(f)(-x)$;

(ii) commutes with translation;

(iii) commutes with dilation;

(iv) $\exists A(p)$ *such that* $\|\mathscr{H} f\|_p \leq A(p)\|f\|_p$, $\forall f \in L^p(\mathbb{R})$. *This result is called the theorem of M. Riesz;*

(v) is uniquely defined by properties (i)-(iv) modulo a multiplicative real constant;

(vi) the corresponding Fourier multiplier is given by $\widehat{\mathscr{H} f}(x) = -i\,\mathrm{sgn}(x)\hat{f}(x)$;

(vii) $\mathscr{H}^2(f) = -f$.

Proof. ad (i)

$$\mathscr{H}(f(-\,\cdot\,))(x) = \frac{1}{\pi}\lim_{\epsilon\to 0}\int_{|t|>\epsilon} \frac{f(t-x)}{t} dt = \frac{1}{\pi}\lim_{\epsilon\to 0}\int_{|t|>\epsilon} -\frac{f(-x-t)}{t} dt = -\mathscr{H}(f)(-x).$$

ad (ii)

$$\mathscr{H}(T_l f)(x) = \frac{1}{\pi}\lim_{\epsilon\to 0}\int_{|t|>\epsilon} \frac{f(x-l-t)}{t} dt = \mathscr{H}(f)(x-l) = T_l\mathscr{H}(f)(x).$$

ad (iii)

$$\begin{aligned}\mathscr{H}(D_s f)(x) &= s^{-1/2}\frac{1}{\pi}\lim_{\epsilon\to 0}\int_{|t|>\epsilon} \frac{f(s(x-t))}{t} dt = s^{-1/2}\frac{1}{\pi}\lim_{\epsilon\to 0}\int_{|t|>\epsilon} \frac{f(sx-\tau)}{\tau/s}\frac{1}{s} d\tau \\ &= D_s\mathscr{H}(f)(x).\end{aligned}$$

ad (iv) We will not proof this here. For a proof see [59], Chapter 3.

ad (v) We will proof this in a more general case in Theorem 2.1.2.

ad (vi) This will be proven in Theorem 2.1.2 in a more general setting.

ad (vii) For the Fourier transforms it is obviously true that

$$\widehat{\mathcal{H}^2(f)}(t) = -i\,\mathrm{sgn}(t)(-i\,\mathrm{sgn}(t))\widehat{f}(t) = -\widehat{f}(t).$$

□

1.4.2 Analytical signals

Definition 1.4.3 (The analytical signal)
LET $f \in L^p(\mathbb{R},\mathbb{R})$, $1 < p < \infty$ *be a 1-D real-valued signal.*
THEN *the* **analytical signal** f_a *is defined by*

$$f_a := f + i\mathcal{H}f \in L^p(\mathbb{R},\mathbb{C}).$$

Theorem 1.4.4 (Properties of analytical signals)
LET $f \in L^p$, $p \in]1,\infty[$.
THEN

(i) $\mathcal{F}(f_a)(w) := \mathcal{F}(f)(w) + \mathrm{sgn}(w)\mathcal{F}(f)(w)$, *i.e., the Fourier transform of the analytical signal vanishes for negative frequencies.*

(ii) $f_a \in L^p(\mathbb{R},\mathbb{C})$ *and* $f = Re(f_a)$, *i.e., the analytical signal is complex-valued and the real-valued signal is its real part.*

(iii) An analytical signal can be decomposed into $f_a(t) = |f_a(t)|e^{i\phi(t)}$, *where* $a(t) = |f_a(t)|$ *is called the* **amplitude** *and* $\phi(t) : \mathbb{R} \to [0,2\pi[$ *the* **phase**, *which is uniquely defined by this decomposition.*

Often the phase defined by the above decomposition is interpreted as the phase of a local Fourier transform assuming that the behaviour under translation, modulation, and rotation is the same as for the Fourier transform.

The use of the phase is inspired by a paper of Oppenheim and Lim [35], who claim that the important information of an image is in the phase of the Fourier transform.

However, it should be noted that this interpretation for the phase is somewhat questionable since the phase acquired by the analytical signal is at best approximately equal to the phase of a local Fourier transform and Oppenheim and Lim argue in the same paper that their arguments do not hold for a localised Fourier transform.

The above decomposition of the analytical signal implies that the signal itself may be uniquely decomposed into amplitude and frequency part via

$$f(t) = a(t)\cos(\phi(t)), \tag{1.1}$$

where $a(t)$ is the amplitude of the analytical signal and $\phi(t)$ is just the phase of the analytical signal.

Inspired by this decomposition is the definition of the analytic instanteous frequency of a signal.

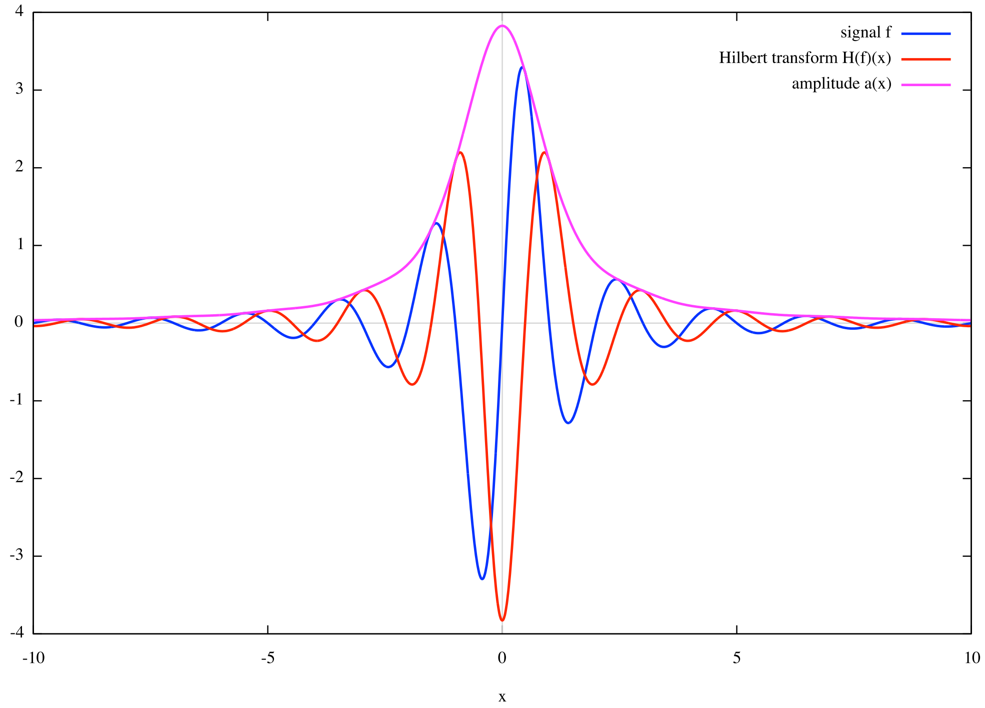


Figure 1.3: Example of an analytical signal: $f(x) = 4\frac{\sin(\pi x)}{x^2+1}$, $\mathscr{H}f(x) = 4\frac{e^{-\pi}-\cos(\pi x)}{x^2+1}$. This example can be found in [30].

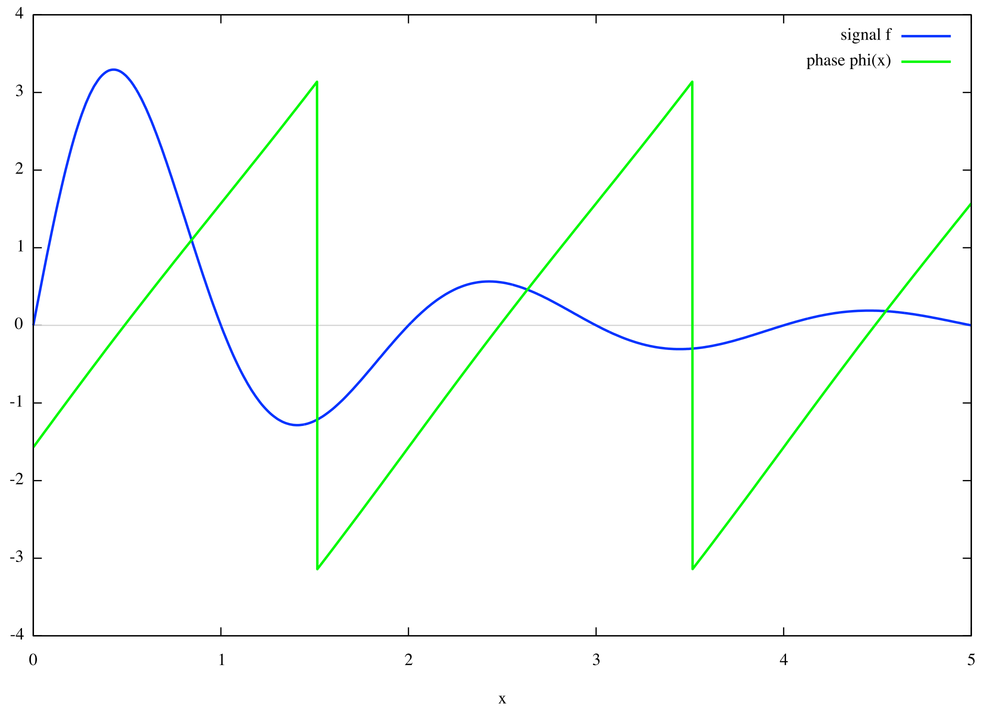


Figure 1.4: Phase of the function f.

Definition 1.4.5 (Instantaneous Frequency)
LET $f \in L^p(\mathbb{R})$, *where* $1 < p < \infty$.
THEN *the* **instantaneous frequency** $\omega(t)$ *of a signal* f *is defined on its support by*

$$\omega(t) := \begin{cases} \dfrac{\mathscr{H}(f)(t)f'(t) - f(t)\mathscr{H}(f)'(t)}{a^2(t)}, & \forall t : a(t) \neq 0; \\ 0, & otherwise. \end{cases}$$

Because

$$\frac{\mathscr{H}(f)(t)}{f(t)} = \frac{a(t)\sin(\phi(t))}{a(t)\cos(\phi(t))} = \tan(\phi(t))$$

and

$$\begin{aligned} \phi'(t) &= \frac{d}{dt}\arctan\left(\frac{\mathscr{H}(f)(t)}{f(t)}\right) = \frac{1}{1+\left(\frac{\mathscr{H}(f)(t)}{f(t)}\right)^2}\frac{\mathscr{H}(f)(t)f'(t) - f(t)\mathscr{H}(f)'(t)}{f^2(t)} \\ &= \frac{f^2(t)}{f^2(t)+\mathscr{H}(f)^2(t)}\frac{\mathscr{H}(f)(t)f'(t) - f(t)\mathscr{H}(f)'(t)}{f^2(t)} = \frac{\mathscr{H}(f)(t)f'(t) - f(t)\mathscr{H}(f)'(t)}{a^2(t)(\cos^2(\phi(t))+\sin^2(\phi(t))} \\ &= \frac{\mathscr{H}(f)(t)f'(t) - f(t)\mathscr{H}(f)'(t)}{a^2(t)} \end{aligned}$$

This simply equals

$$\omega(t) = \phi'(t),$$

where $\phi(t)$ is the phase obtained by the decomposition (1.1) of the signal f.

Chapter 2

Riesz transforms

This chapter will introduce the Riesz transform. The first part of this chapter will state some classical results on the Riesz transform, give an interpretation for the term "analytical" in the name "analytical signal" and state that the Riesz transforms are steerable.

The second part of this chapter states new results. We will define the Riesz transform of two classes of distributions – distributions modulo polynomials (Definition 2.2.10) and tempered distributions whose support in the Fourier domain does not contain the origin (Theorem 2.2.13). As applications by means of a certain new operator between $\mathscr{S}'(\mathbb{R})$ and $\mathscr{S}'(\mathbb{R}^n)$ (Proposition 2.2.14) we state a novel connection between the Hilbert and the Riesz transform (Theorem 2.2.15) and give a novel Bedrosian identity for Hilbert transforms of distributions (Theorem 2.2.18) and Riesz transforms of distributions (Theorem 2.2.19).

In the following let $\mathbb{R}^{n+1}_+$ denote the upper half space

$$\mathbb{R}^{n+1}_+ := \{(x, x_0) : x \in \mathbb{R}^n,\ x_0 > 0\}$$

for which $\mathbb{R}^n$ is the boundary hyperplane. Let $\{e_\alpha\}_{\alpha=1}^n \subset \mathbb{R}^{n+1}_+$ be an orthonormal basis of $\mathbb{R}^n \subset \mathbb{R}^{n+1}_+$ and let $e_0 \in \mathbb{R}^{n+1}_+$ such that $\{e_\alpha\}_{\alpha=0}^n$ is a orthonormal basis of $\mathbb{R}^{n+1}$.

2.1 Riesz transforms and conjugate harmonic functions

2.1.1 Riesz transforms

This section will introduce the Riesz transform as the unique extension of the Hilbert transform to higher dimensions. I.e. the Riesz transform is the only transform satisfying certain reasonable properties that reduces to the Hilbert transform in the one dimensional case. (See Theorem 2.1.2 and Theorem 2.1.7.)

Furthermore we will note the correspondence between the term analytical in the expression analytical signal and the term analytical function.

Finally we will state in which respect the Riesz transform is steerable.

Definition 2.1.1 (The partial Riesz transform)
LET $f \in L^p(\mathbb{R}^n)$; $1 < p < \infty$.

THEN *the* **partial Riesz transforms** *are defined by*

$$R_\alpha : L^p(\mathbb{R}^n,\mathbb{R}) \to L^p(\mathbb{R}^n,\mathbb{R}),\ R_\alpha f(x) := \lim_{\delta\to 0} \frac{\Gamma(\frac{n+1}{2})}{\pi^{(n+1)/2}} \int_{|y|>\delta} \frac{y_\alpha}{|y|^{n+1}} f(x-y)\,dy,$$

where $\alpha \in \{1,\dots,n\}$.
The **Riesz transform** *is the operator*

$$R : L^p(\mathbb{R}^n,\mathbb{R}) \to L^p(\mathbb{R}^n,\mathbb{R}^n),\ f \mapsto (R_1 f,\dots,R_n f).$$

In the following we will refer to dilations of the form $D_t f(x) = |t|^{-\frac{n}{2}} f(tx)\ \forall t > 0,\ x \in \mathbb{R}^n$, as isotropic dilations.

Theorem 2.1.2 (Properties of the Riesz transformations)
LET $\alpha \in \{1,\dots,n\}$ *and* $f \in L^p(\mathbb{R}^n)$, *where* $1 < p < \infty$.
THEN *the following properties of the Riesz transforms hold:*

(i) LET $\rho \in O(n)$ *be a rotation or a reflection which acts on functions by* $\rho(f)(x) := f(\rho^{-1}x)$.

THEN

$$\rho^{-1} R_\alpha \rho f = \sum_{\beta=1}^{n} \rho_{\alpha\beta} R_\beta f;$$

(ii) R_α *commutes with translation;*

(iii) R_α *commutes with isotropic dilations;*

(iv) $\exists A(p,n)$ *such that* $\|R_\alpha f\|_p \le A(p,n)\|f\|_p,\ \forall f \in L^p(\mathbb{R}^n)$;

(v) R_α *is uniquely defined by properties (i)-(iv) up to a multiplicative complex constant;*

(vi) The Fourier multiplier of the Riesz transform is $\widehat{R_\alpha f}(x) = i\frac{x_\alpha}{|x|}\widehat{f}(x)$;

(vii) $\sum_{\alpha=1}^n R_\alpha^2(f) = -f$;

(viii) $R_\alpha = -\frac{\partial}{\partial x_\alpha}(-\triangle)^{-\frac{1}{2}}$.

Proof. The Proof will follow along the lines of [45] chapter III.

ad (i)

$$\begin{aligned}
\rho^{-1}R_\alpha\rho f(x) &= \rho^{-1}R_\alpha f(\rho^{-1}x)\\
&= \rho^{-1}\lim_{\delta\to 0}\frac{\Gamma(\frac{n+1}{2})}{\pi^{(n+1)/2}}\int_{|y|>\delta}\frac{y_\alpha}{|y|^{n+1}}f(\rho^{-1}(x-y))dy\\
&= \lim_{\delta\to 0}\frac{\Gamma(\frac{n+1}{2})}{\pi^{(n+1)/2}}\int_{|y|>\delta}\frac{y_\alpha}{|y|^{n+1}}f(x-\rho^{-1}y)dy\\
&= \lim_{\delta\to 0}\frac{\Gamma(\frac{n+1}{2})}{\pi^{(n+1)/2}}\int_{|y|>\delta}\frac{\rho y_\alpha}{|\rho y|^{n+1}}f(x-y)dy\\
&= \lim_{\delta\to 0}\frac{\Gamma(\frac{n+1}{2})}{\pi^{(n+1)/2}}\int_{|y|>\delta}\frac{\sum_{\beta=1}^n\rho_{\alpha\beta}y_\beta}{|y|^{n+1}}f(x-y)dy\\
&= \sum_{\beta=1}^n\rho_{\alpha\beta}\lim_{\delta\to 0}\frac{\Gamma(\frac{n+1}{2})}{\pi^{(n+1)/2}}\int_{|y|>\delta}\frac{y_\beta}{|y|^{n+1}}f(x-y)dy\\
&= \sum_\beta\rho_{\alpha\beta}R_\beta f(x).
\end{aligned}$$

ad (ii)

$$\begin{aligned}
R_\alpha(T_lf)(x) &= \frac{\Gamma(\frac{n+1}{2})}{\pi^{(n+1)/2}}\lim_{\epsilon\to 0}\int_{|y|>\epsilon}\frac{y_\alpha f(x-l-y)}{|y|^{n+1}}dy = R_\alpha(f)(x-l)\\
&= T_lR_\alpha(f)(x).
\end{aligned}$$

ad (iii) Let $D_t: f(x)\to|t|^{-\frac{n}{2}}f(tx)$ denote the dilation operator, then

$$\begin{aligned}
R_\alpha D_tf(x) &= |t|^{-\frac{n+1}{2}}R_\alpha f(tx)\\
&= |t|^{-\frac{n}{2}}\lim_{\delta\to 0}\frac{\Gamma(\frac{n+1}{2})}{\pi^{(n+1)/2}}\int_{|y|>\delta}\frac{y_\alpha}{|y|^{n+1}}f(t(x-y))dy\\
&= |t|^{-\frac{n}{2}}\lim_{\delta\to 0}\frac{\Gamma(\frac{n+1}{2})}{\pi^{(n+1)/2}}\int_{|y|>t\delta}\frac{t^{-1}z_\alpha}{|t^{-1}z|^{n+1}}f(tx-z)t^{-n}dz\\
&= |t|^{-\frac{n}{2}}\lim_{\delta\to 0}\frac{\Gamma(\frac{n+1}{2})}{\pi^{(n+1)/2}}\int_{|y|>t\delta}\frac{z_\alpha}{|z|^{n+1}}f(tx-z)dz\\
&= D_tR_\alpha f(x).
\end{aligned}$$

ad (iv) A proof can be found in [45] II.4.2, Theorem 3.

ad (v) First we will proof the following

Lemma 2.1.3
LET $m:\mathbb{R}^n\to\mathbb{R}^n$ *be homogeneous of degree* 0*, i.e.,* $m(tx)=m(x),\ \forall x\in\mathbb{R}^n,\ t\in\mathbb{R}_+$, *and let*

$$m(\rho x)=\rho(m(x)). \tag{2.1}$$

THEN

$$\exists c\in\mathbb{R}: m(x)=c\frac{x}{|x|}.$$

Proof of the Lemma. It suffices to proof the Lemma in the case that $x \in S^{n-1}$. Let $e_1, \dots, e_n$ be the orthonormal basis of $\mathbb{R}^n$, let $c = m_1(e_1)$ and let ϱ be any rotation of $\mathbb{R}^n$ leaving e_1 fixed. Then by (2.1) it follows that

$$m_\alpha(e_1) = \sum_\beta \varrho_{\alpha\beta} m_\beta(e_1)$$

and hence the $(n-1)$-dimensional vector $(m_2(e_1), \dots, m_n(e_1))^T$ is invariant under all $(n-1)$-dimensional rotations. Thus $m_2(e_1) = \dots = m_n(e_1) = 0$.

It follows that $m_\alpha(\rho e_1) = \rho_{\alpha 1} m_1(e_1) = c\rho_{\alpha 1}$ for any rotation ρ of $\mathbb{R}^n$. Now let $x = \rho e_1$, then $\rho_{\alpha 1} e_1 = x_\alpha$ and thus $m_\alpha(x) = c x_\alpha$. □

Let now $\{T_\alpha\}_{\alpha=1}^n$ be a family of operators which satisfy (i)-(iv). Then by (ii) and (iv) it follows that the T_α can be realized by bounded Fourier multipliers $\widehat{T_\alpha} = m_\alpha$.

Now (iii) shows that

$$\begin{aligned} m_\alpha(tx)\widehat{f}(x) &= t^{-\frac{n}{2}} D_t(m_\alpha(x)\widehat{f}(\frac{x}{t})) = D_t(m_\alpha(x) D_{t^{-1}}\widehat{f}(x)) \\ &= \mathscr{F}(D_{-t} T_\alpha D_t f)(x) = \mathscr{F}(T_\alpha f)(x) \\ &= m_\alpha(x) f(x), \end{aligned}$$

i.e. m is homogeneous of degree 0.

Using (i) we see that on the one hand

$$\begin{aligned} \mathscr{F}(\rho^{-1} T_\alpha \rho f)(x) &= \rho^{-1} m_\alpha(x)(\widehat{\rho f})(x) = \rho^{-1}(m_\alpha(x)\widehat{f}(\rho^{-1}x)) \\ &= m_\alpha(\rho x) f(x) \end{aligned}$$

and on the other hand

$$\mathscr{F}\Big(\sum_\beta \rho_{\alpha\beta} T_\beta f\Big)(x) = \sum_\beta \rho_{\alpha\beta} \mathscr{F} T_\beta f(x) = \sum \rho_{\alpha\beta} m_\beta(x)\widehat{f}(x).$$

Thus the m_α suffice (2.1) and thus our Lemma, can be applied. Hence $m_\alpha = c\frac{x_\alpha}{|x|}$ for some constant c.

ad (vii) To calculate the Fourier multiplier corresponding to the Riesz transform we use the fact that $\frac{\Gamma(\frac{n+1}{2})}{\pi^{(n+1)/2}} x_\alpha$ is a homogeneous harmonic polynomial of degree 1 and Theorem 5 in chapter III of [45]:

Lemma 2.1.4
LET $P_\beta(x)$ *be a homogeneous harmonic polynomial of degree* β, $\beta \geq 1$.

THEN *the Fourier multiplier corresponding to the transformation*

$$Tf(x) = lim_{\epsilon\to 0} \int_{|y|\geq\epsilon} \frac{\Omega(y)}{|y|^n} f(x-y)dy$$

with $\Omega(x) = \frac{P_\beta(x)}{|x|^\beta}$ *is* $\gamma_\beta \frac{P_\beta(x)}{|x|^\beta}$, *where* $\gamma_\beta = i^\beta \pi^{n/2} \frac{\Gamma(\frac{\beta}{2})}{\Gamma(\frac{\beta+n}{2})}$

Using this Lemma we get $m_\alpha = \frac{\Gamma(\frac{n+1}{2})}{\pi^{(n+1)/2}} \gamma_1 \frac{x_\alpha}{|x|} = i\frac{x_\alpha}{|x|}$.

ad (vi) For the Fourier multipliers (see (vi)) it obviously holds true that

$$\mathscr{F}(\sum_{\alpha=1}^{n} R_\alpha^2(f))(t) = \begin{cases} (-i)(-i)\sum_{\alpha=1}^{n} \frac{t_\alpha^2}{|t|}\widehat{f}(t) = -\widehat{f}(t) & \forall t \neq 0 \\ 0 & \text{if } t = 0 \end{cases}$$

ad (vii) This can easily be seen by computing the corresponding Fourier multipliers.

$$\mathscr{F}\left(R_\alpha f\right)(x) = i\frac{x_\alpha}{|x|}\widehat{f}(x) = 2\pi i x_\alpha(-i^2 4\pi^2|x|^2)^{-1/2}\widehat{f}(x) = \mathscr{F}\left(-\frac{\partial}{\partial_{\cdot\alpha}}(-\triangle)^{-1/2} f\right)(x).$$

□

2.1.2 Conjugate harmonic functions

Conjugate harmonic functions are the connection between the term analytical in the expression analytical signal and the term analytical function. Furthermore they add a further respect in which the Riesz transform is the unique extension of the Riesz transform.

In order to define conjugate harmonic functions we will first need to introduce the Poisson transform:

Definition 2.1.5 (The Poisson transform)
LET $x_0 > 0$.
THEN *the* **Poisson** *kernel* P_{x_0} *on* $\mathbb{R}^n$ *is defined by*

$$P_{x_0}(x) := \int_{\mathbb{R}^n} e^{2\pi i\langle t,x\rangle} e^{-2\pi|t|x_0} dt.$$

LET $f \in L^p(\mathbb{R}^n)$, $1 < p < \infty$.
THEN *the Poisson kernel defines via convolution the* **Poisson transform** u_f

$$u_f(x, x_0) := (P_{x_0} * f)(x).$$

The Poisson transform maps a function in $L^p(\mathbb{R}^n)$ to a function in $L^p(\mathbb{R}^{n+1})$.

Theorem 2.1.6 (Properties of the Poisson transform)
LET $f \in L^p(\mathbb{R}^n)$, $1 < p < \infty$, *and let* $u_f(x, x_0) = (P_{x_0} * f)(x)$ *be its Poisson transform.*
THEN

(i) $P_{x_0}(x) = \frac{\Gamma(\frac{n+1}{2})}{\pi^{\frac{n+1}{2}}} \frac{x_0}{(|x|^2+x_0^2)^{\frac{n+1}{2}}}$;

(ii) $\triangle u_f = \frac{\partial^2 u_f}{\partial x_0^2} + \sum_{\alpha=1}^{n} \frac{\partial^2 u_f}{\partial x_\alpha^2} = 0$, *i.e., the Poisson transform of an* L^p*-function is harmonic;*

(iii) $P_{x_0} > 0$;

(iv) $\int_{R^n} P_{x_0}(x)dx = 1$, $\forall x_0 > 0$;

(v) $P_{x_0}(x)$ *is a decreasing function of* $|x|$;

(vi) $P_{x_0}(x)$ *is harmonic in* $\mathbb{R}_+^{n+1}$;

(vii) *$P_{x_0}(x) \in L^p(\mathbb{R}^n)$, $1 \le p \le \infty$, and for $f \in L^p(\mathbb{R}^n)$ it follows that $u_f \in L^p(\mathbb{R}^{n+1})$ is harmonic in R_+^{n+1};*

(viii) *Let $u(x, x_0)$ be harmonic in $\mathbb{R}_+^{n+1}$. Then u is the Poisson transform of a function in $L^p(\mathbb{R}^n)$ if and only if $\sup_{x_0>0} \|u(x, x_0)\|_p < \infty$;*

(ix) *Semi-group property: $P_l * P_r = P_{l+r}\ \forall l, r > 0$;*

(x) *$\lim_{x_0 \to 0} u_f(x, x_0) = f(x)$ for almost every x and $\lim_{x_0 \to 0} \|f(x) - u(x, x_0)\|_p = 0$, $1 \le p < \infty$;*

(xi) *$P_{x_0}(x)$ is homogenous of degree $-n$: $P_a(x) = P_1(x/a)a^{-n}$, $a > 0$;*

(xii) *The Fourier multiplier corresponding to the Poisson transform is $e^{-2\pi|t|x_0}$;*

Proof. See [45] Chapters III.2 and III.4.3. □

The connection between the analytical signal and analytical functions are a set of generalized Cauchy-Riemann equations which the Poisson transforms of the analytical signal respectively the Poisson transforms of the Riesz transforms and the original function obey. This theorem can be found in [45] Theorem 3 in Chapter III.

Theorem 2.1.7 (Conjugate harmonic functions)
LET *f and f_α, $\alpha = 1, \ldots, n$, belong to $L^p(\mathbb{R}^n)$, $1 < p < \infty$, and let*

$$u_0(x, x_0) := P_{x_0} * f(x), \quad u_\alpha(x, x_0) := P_{x_0} * f_\alpha(x).$$

THEN

$$f_\alpha = -R_\alpha(f), \ \alpha = 1, \ldots, n,$$

if and only if the generalized Cauchy-Riemann equations hold:

$$\sum_{\alpha=0}^{n} \frac{\partial u_\alpha}{\partial x_\alpha} = 0, \qquad (2.2)$$
$$\frac{\partial u_\alpha}{\partial x_\beta} = \frac{\partial u_\beta}{\partial x_\alpha};\ \alpha, \beta = 0, \ldots, n.$$

Proof. We will give the proof in the case $p = 2$. The theorem is equivalent to a corresponding theorem in [45] Chapter III.4.4, a reference for the complete proof can be found there. Note however that the Fourier transform is defined differently there.
Suppose $f_\alpha = -R_\alpha f$, then $\widehat{f_\alpha}(t) = \frac{-it_\alpha}{|t|}\widehat{f}(t)$ and, hence,

$$u_\alpha(x, x_0) = \int_{\mathbb{R}^n} \widehat{f}(t) \frac{-it_\alpha}{|t|} e^{-2\pi|t|x_0} e^{2\pi i t x} dt,$$

and

$$u(x, x_0) = \int_{\mathbb{R}^n} \widehat{f}(t) e^{-2\pi|t|x_0} e^{2\pi i t x} dt.$$

Because of the dominated convergence theorem [41], we may differentiate under the integral sign, and hence,

$$\frac{\partial u}{\partial x_0}(x,x_0) = -2\pi \int_{\mathbb{R}^n} \widehat{f}(t)|t|e^{-2\pi|t|x_0}e^{2\pi i t x}dt.$$
$$\frac{\partial u}{\partial x_\beta}(x,x_0) = 2\pi i \int_{\mathbb{R}^n} \widehat{f}(t)t_\beta e^{-2\pi|t|x_0}e^{2\pi i t x}dt.$$
$$\frac{\partial u_\beta}{\partial x_0}(x,x_0) = 2\pi i \int_{\mathbb{R}^n} \widehat{f}(t)t_\beta e^{-2\pi|t|x_0}e^{2\pi i t x}dt.$$
$$\frac{\partial u_\alpha}{\partial x_\beta}(x,x_0) = 2\pi \int_{\mathbb{R}^n} \widehat{f}(t)\frac{t_\alpha t_\beta}{|t|} e^{-2\pi|t|x_0}e^{2\pi i t x}dt.$$

Now the generalized Cauchy-Riemann conditions are easy to check. The first one follows from

$$\begin{aligned}\sum_{\alpha=1}^n \frac{\partial u_\alpha}{\partial x_\alpha}(x,x_0) &= \sum_{\alpha=1}^n 2\pi \int_{\mathbb{R}^n} \widehat{f}(t)\frac{t_\alpha^2}{|t|}e^{-2\pi|t|x_0}e^{2\pi i t x}dt \\ &= 2\pi \int_{\mathbb{R}^n} \widehat{f}(t)\frac{\sum_{\alpha=1}^n t_\alpha^2}{|t|}e^{-2\pi|t|x_0}e^{2\pi i t x}dt \\ &= 2\pi \int_{\mathbb{R}^n} \widehat{f}(t)|t|e^{-2\pi|t|x_0}e^{2\pi i t x}dt \\ &= -\frac{\partial u}{\partial x_0}(x,x_0),\end{aligned}$$

the second equation follows from

$$\begin{aligned}\frac{\partial u_\alpha}{\partial x_\beta}(x,x_0) &= 2\pi \int_{\mathbb{R}^n} \widehat{f}(t)\frac{t_\alpha t_\beta}{|t|}e^{-2\pi|t|x_0}e^{2\pi i t x}dt \\ &= \frac{\partial u_\beta}{\partial x_\alpha}(x,x_0),\end{aligned}$$

and

$$\begin{aligned}\frac{\partial u}{\partial x_\beta}(x,x_0) &= 2\pi i \int_{\mathbb{R}^n} \widehat{f}(t)t_\beta e^{-2\pi|t|x_0}e^{2\pi i t x}dt \\ &= \frac{\partial u_\beta}{\partial x_0}(x,x_0).\end{aligned}$$

Conversely, let $\alpha \in \{1,\dots,n\}$ and $u_\alpha(x,x_0) = \int_{\mathbb{R}^n} \widehat{f_\alpha}(t)e^{-2\pi|t|x_0}e^{2\pi i t x}dt$. The fact that $\frac{\partial u_0}{\partial x_\alpha} = \frac{\partial u_\alpha}{\partial x_0}$, shows that

$$2\pi i \int_{\mathbb{R}^n} \widehat{f}(t)t_\alpha e^{-2\pi|t|x_0}e^{2\pi i t x}dt = -2\pi \int_{\mathbb{R}^n} \widehat{f_\alpha}(t)|t|e^{-2\pi|t|x_0}e^{2\pi i t x}dt.$$

Therefore by the uniqueness of the inverse Fourier transform $\widehat{f_\alpha}(t) = -\frac{i t_\alpha}{|t|}\widehat{f}(t)$, and so

$$f_\alpha = -R_\alpha(f),\ \alpha = 1,\dots,n.$$

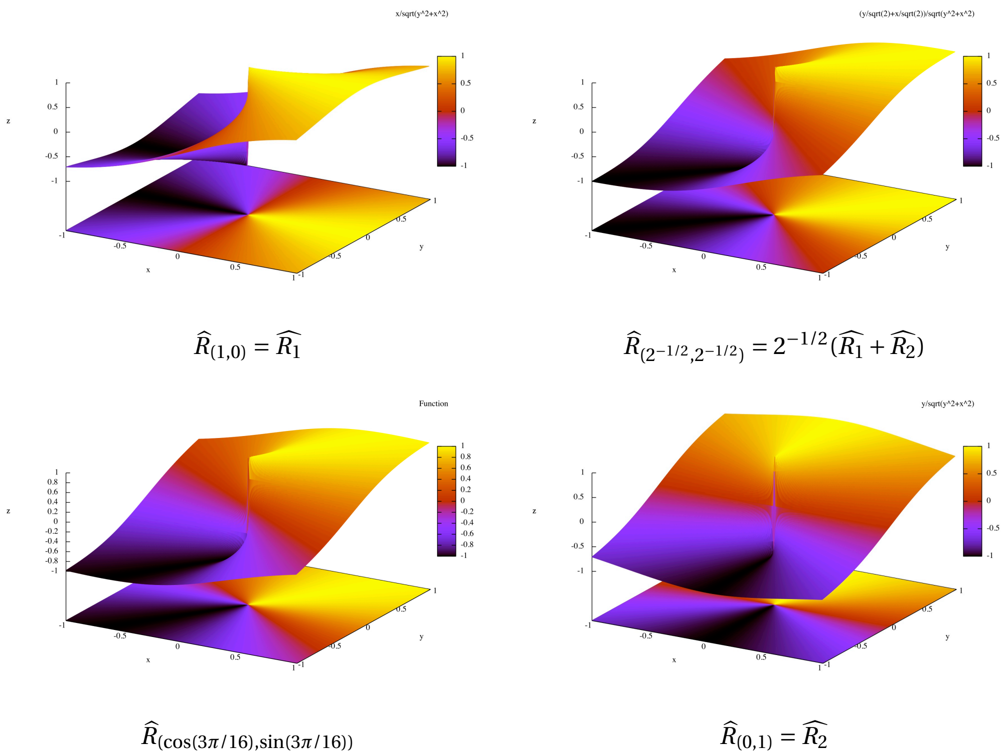
□

Remark 2.1.8
The equations (2.2) *are not fully satisfactory because the analytical signal satisfying this equation would be* $\begin{pmatrix} f \\ -\mathscr{H}f \end{pmatrix}$ *– that is a vector valued function rather than* $f_a = f + i\mathscr{H}f$ *the complex valued monogenic signal.*

The reason for this difference is that equations (2.2) *do not reduce to the Cauchy-Riemann equations of complex analysis.*

To fix this in Theorem 4.5.2 we will define the analog of the analytical signal for Riesz transforms – the monogenic signal in chapter 4 and the analog of complex analysis for higher dimensions – Clifford analysis in section 3.3.

2.1.3 Steerabilty of the Riesz transform



$\widehat{R}_{(1,0)} = \widehat{R_1}$ $\widehat{R}_{(2^{-1/2},2^{-1/2})} = 2^{-1/2}(\widehat{R_1} + \widehat{R_2})$

$\widehat{R}_{(\cos(3\pi/16),\sin(3\pi/16))}$ $\widehat{R}_{(0,1)} = \widehat{R_2}$

Figure 2.1: Steerability of the Fourier multiplier of the Riesz transform: In the upper left and in the lower right corner the Fourier multipliers of the Riesz transforms with respect to the basis directions are given, the other two images portrait linear combinations - the Fourier multipliers with respect to the directions corresponding to the linear combination of the basis directions.

LET $E = \{e_1, \ldots, e_n\}$ and $D = \{d_1, \ldots, d_n\}$ be two ONBs of $\mathbb{R}^n$. A linear orthogonal mapping $\rho \in O(\mathbb{R}^n)$ is then uniquely defined by $d_l = \rho e_l$, $\forall l = 1, \ldots, n$.

THEN according to Theorem 2.1.2[i] the Riesz transforms with respect to the basis D can be expressed via the Riesz transforms with respect to the basis E as

$$R_{d_l} f(x) = \sum_{\beta=1}^{n} \rho_{l,\beta} R_{e_\beta}(x).$$

As a consequence the Riesz transform with respect to an arbitrary direction is well defined as a linear combination of the Riesz transforms with respect to the basis directions.

Definition 2.1.9 (Steerability of the Riesz transform)
LET $d \in \mathbb{R}^n : |d| = 1$ *and let* $\rho \in SO(\mathbb{R}^n) : d = \rho e_1$.
THEN *the Riesz transform in direction* d *is given by*

$$R_d f(x) = \rho^{-1} R_1 \rho f(x) = \sum_{\beta=1}^{n} \rho_{1,\beta} R_\beta f(x),\ \forall f \in L^2(\mathbb{R}^n),\ x \in \mathbb{R}^n. \tag{2.3}$$

From this definition it is clear, that the Riesz transform is steerable, since the Riesz transform with respect to any direction is a linear combination of the n Riesz transforms with respect to the basis directions. This is illustrated in Figure 2.1.

2.2 Riesz transforms of distributions

In this section we will define the Riesz transform of a distribution. To do this we will consider the Lizorkin space, a subspace of the Schwartz space which is invariant under the Riesz transform. Then we will define the Riesz transform for elements in the dual of this space, which is the space of tempered distributions modulo polynomials and for a certain class of tempered distributions.

The Riesz transform on the dual of the Lizorkin space we define seems to be new as well as the Riesz transform on the subset of tempered distributions we will consider. Since the tempered distributions we consider are not necessarily compactly supported they complement the known theory of Riesz transforms of compactly supported distributions. The Hilbert transform of a compactly supported distribution is easily defined via convolution with the Cauchy principal value distribution $\mathrm{PV}\int_{\mathbb{R}} \frac{\phi(x)}{x} dx,\ \forall \phi \in \mathfrak{D}(\mathbb{R}^n)$. (See [30] for details on the Hilbert transforms of distributions.)

In subsection 2.2.6 it will become apparent that our definition of the Riesz transform of distributions is especially well suited to proof a novel Bedrosian identity for distributions. Indeed the prerequisites for the Bedrosian identity are all that is required for the Hilbert transform of the involved distributions to be well defined. Furthermore we will give the connection between the Riesz transform of certain distributions and the Hilbert transform. This connection will then yield a Bedrosian identity for Riesz transforms of certain distributions.

2.2.1 Preliminaries: Lizorkin spaces and derivatives of the Riesz multiplier

Definition 2.2.1 (Lizorkin spaces)
The **Lizorkin space** $\Phi(\mathbb{R}^n)$ *is the linear subspace of* $\mathscr{S}(\mathbb{R}^n)$ *defined as*

$$\Phi(\mathbb{R}^n) = \left\{ \phi \in \mathscr{S}(\mathbb{R}^n) : \int_{\mathbb{R}^n} x^\gamma \phi(x) dx = 0,\ \forall \gamma = \{\gamma_1, \dots, \gamma_n\} \in \mathbb{N}_0^n \right\}.$$

Thus the Lizorkin space is the subspace of the Schwartz space that consists of functions with arbitrarily many vanishing moments.

Proposition 2.2.2 (Lizorkin spaces and distributions)
LET $\Phi(\mathbb{R}^n)$ *be the Lizorkin space.*

THEN $\Phi(\mathbb{R}^n)$ *is a closed subspace of* $\mathscr{S}(\mathbb{R}^n)$.

LET $\Psi(\mathbb{R}^n) = \left\{\psi \in \mathscr{S}(\mathbb{R}^n) : \partial^\gamma \psi(0) = 0,\ \forall \gamma = \{\gamma_1, \ldots, \gamma_n\} \in \mathbb{N}_0^n\right\}$.

THEN $\Phi(\mathbb{R}^n)$ *and* $\Psi(\mathbb{R}^n)$ *are a Fourier pair, i.e.,* $\mathscr{F} : \Phi(\mathbb{R}^n) \to \Psi(\mathbb{R}^n)$ *and* $\mathscr{F} : \Psi(\mathbb{R}^n) \to \Phi(\mathbb{R}^n)$. *The topologies on* $\Phi(\mathbb{R}^n) \subset \mathscr{S}(\mathbb{R}^n)$ *and* $\Psi(\mathbb{R}^n) \subset \mathscr{S}(\mathbb{R}^n)$ *are the subspace topologies. The topological dual* $\Phi'(\mathbb{R}^n)$ *is the space of* **tempered distributions modulo polynomials**.

Proof. See [31]. □

Remark 2.2.3
The Lizorkin space $\Phi(\mathbb{R}^n)$ *is often denoted as* $\mathscr{S}_0(\mathbb{R}^n)$ *[25] or even* $\mathscr{S}_\infty(\mathbb{R}^n)$.

Definition 2.2.4 (The space $\mathscr{E}_0(\mathbb{R}^n)$)
Let us denote $\mathscr{E}_0(\mathbb{R}^n) := \{f \in C^\infty(\mathbb{R}^n) : \partial^\gamma f(0) = 0,\ \forall \gamma \in \mathbb{N}_0^n\}$. *Obviously* $\Psi(\mathbb{R}^n) = \mathscr{S}(\mathbb{R}^n) \cap \mathscr{E}_0(\mathbb{R}^n)$.

We will use the following characterization of $\mathscr{E}_0$ from [31].

Lemma 2.2.5 (Characterizations of $\mathscr{E}_0$)
LET $f \in C^\infty(\mathbb{R}^n)$.

THEN *the following are equivalent:*

1. $f \in \mathscr{E}_0(\mathbb{R}^n)$.
2. $\lim_{|x|\to 0} \frac{\partial^\gamma f(x)}{|x|^l} = 0;\ \forall \gamma \in \mathbb{N}_0^n,\ l \in \mathbb{N}_0$.
3. $\lim_{|x|\to 0} \frac{f(x)}{|x|^l} = 0;\ \forall l \in \mathbb{N}_0$.

The above Lemma implies that to proof that the Riesz transform maps the Lizorkin space $\Phi(\mathbb{R}^n)$ into itself, we have to take a look at the partial derivatives of the Fourier multiplier

$$\widehat{R}_\alpha : \mathbb{R}^n \to \mathbb{R},\ \xi \mapsto \frac{i\xi_\alpha}{|\xi|}$$

of the partial Riesz transform.

2.2.2 Derivatives of the Fourier multiplier of the Riesz transform

Lemma 2.2.6 (Partial derivatives of $\widehat{R}_\alpha$)
LET $\beta = (\beta_1, \ldots, \beta_n) \in \mathbb{N}_0^n$ *be a multi index,* $\alpha \in \{1, \ldots, n\}$, $\xi \in \mathbb{R}^n \setminus \{0\}$.

THEN

$$\partial^\beta \frac{\xi_\alpha}{|\xi|} = \beta_\alpha \Theta(\xi; \beta_1, \ldots, \beta_{\alpha-1}, \beta_\alpha - 1, \beta_{\alpha+1}, \ldots, \beta_n) + \frac{\beta_\alpha!}{2(\beta_\alpha - 2)!} \xi_\alpha \Theta(\xi, \beta_1, \ldots, \beta_n). \tag{2.4}$$

Here,

$$\begin{aligned}\Theta(\xi;\beta_1,\dots,\beta_n) &= \frac{(-1)^{\beta}\sum_{j=1}^{n}\frac{1}{\beta_j!}}{|\xi|^{2|\beta|+1}}\xi^{\beta} \\ &+ \sum_{\substack{k_1(\beta_1),k_2(\beta_1)\in T_{\beta_1}\\ \vdots \\ k_1(\beta_n),k_2(\beta_n)\in T_{\beta_n}}} \frac{\beta!}{k_1(\beta)!k_2(\beta)!}\sum_{j=1}^{n}\frac{1}{\big(k_1(\beta_j)+k_2(\beta_j)\big)!}\frac{(-1)^{k_1(\beta)}(-1)^{k_2(\beta)}}{|\xi|^{2(k_1(\beta)+k_2(\beta))+1}}\xi^{k_1(\beta)}\end{aligned}$$

As usual $\beta! = \prod_{j=1}^{n}\beta_j!$ *and* $a^{|\beta|} = a^{\sum_{j=1}^{n}\beta_j}$; $\forall a\in\mathbb{R}$.

T_{β_j} denotes the set of pairs of integers $(k_1,k_2)\in\mathbb{N}_0^2$ such that $\beta_j = k_1(\beta_j)+2k_2(\beta_j)$ and $k_l(\beta) = \big(k_l(\beta_1),\dots,k_l(\beta_n)\big)$.

Proof. For $\alpha\in\{1,\dots,n\}$ and $\xi\in\mathbb{R}^n\setminus\{0\}$ we can write

$$\frac{\xi_\alpha}{|\xi|} = \frac{\xi_\alpha}{\sqrt{\sum_{j=1}^{n}\xi_j^2}} = f(\xi_\alpha)g(h(\xi)),$$

where $f(a)=a$, $g(a)=a^{-\frac{1}{2}}$ $\forall a\in\mathbb{R}$ and $h(\xi)=\sum_{k=1}^{n}\xi_k^2$, $\forall\xi\in\mathbb{R}^n$. Using the product rule on the α-th partial derivative we get

$$\begin{aligned}\frac{\partial^{\beta_\alpha}}{\partial\xi_\alpha^{\beta_\alpha}}f(\xi_\alpha)g(h(\xi)) &= \sum_{k=0}^{\beta_\alpha}\binom{\beta_\alpha}{k}\frac{\partial^{\beta_\alpha-k}}{\partial\xi_\alpha^{\beta_\alpha-k}}f(\xi_j)\frac{\partial^k}{\partial\xi_\alpha^k}g(h(\xi)) \\ &= \beta_\alpha\frac{\partial^{\beta_\alpha-1}}{\partial\xi_\alpha^{\beta_\alpha-1}}g(h(\xi)) + \frac{\beta_\alpha!}{2(\beta_\alpha-2)!}\xi_\alpha\frac{\partial^{\beta_\alpha}}{\partial\xi_\alpha^{\beta_\alpha}}g(h(\xi)).\end{aligned}$$

Now we apply the remaining partial derivatives and set

$$\Theta(\xi;\beta_1,\dots,\beta_n) = \frac{\partial^{\sum_{k=1}^{n}\beta_k}g(h(\xi))}{\prod_{k=0}^{n}\partial\xi_k^{\beta_k}}.$$

The formula of Faá di Bruno Lemma A.1.4 which yields (2.4).

Since in our case $h(\xi)=c+\xi^2$, where $c\in\mathbb{R}$ is independent of ξ, only the subset of T_m, where only coefficients b_1 and b_2 are non-zero, is relevant.

Now the statement follows since $g^{(m)}(\xi) = c\xi^{-\frac{1}{2}-m}$ and $h'(\xi)=2\xi$ and hence $h^{(2)}(\xi)=2$. The constants in the statement follow since the constant factors 2 in the derivatives of h cancel the factors of $\frac{1}{2}$ in the derivatives of g. □

Remark 2.2.7
From the above Lemma it follows that the partial derivatives of $\widehat{R}_\alpha$ *have the form*

$$\partial^\beta\frac{\xi_\alpha}{|\xi|} = \frac{p_{\alpha,\beta}(\xi)}{q_{2|\beta|+1}(|\xi|)},\ \forall\xi\in\mathbb{R}^n\setminus\{0\}, \tag{2.5}$$

where

$$p_{\alpha,\beta}(\xi) = \sum_{k_1=0}^{\beta_1}\cdots\sum_{k_\alpha=0}^{\beta_\alpha+1}\cdots\sum_{k_n=0}^{\beta_n}a_k\xi^k,$$

and $a_k\in\mathbb{C}$, $k=(k_1,\dots,k_n)$ *and* $q_{2|\beta|+1}(|\xi|) = \sum_{j=0}^{2|\beta|+1}b_j|\xi|^j$, *and* $b_j\in\mathbb{C}$.

Theorem 2.2.8 (Partial derivatives of $\widehat{R}_\alpha$ are homogenous)
The partial derivatives of $\widehat{R}_\alpha$ have the form

$$\partial^{\boldsymbol{\beta}} \frac{\xi_\alpha}{|\xi|} = \frac{h_{\alpha,\boldsymbol{\beta}}(\xi)}{|\xi|^{|\boldsymbol{\beta}|}}, \ \forall \xi \in \mathbb{R}^n \setminus \{0\}, \tag{2.6}$$

where $h_{\alpha,\boldsymbol{\beta}} : \mathbb{R}^n \to \mathbb{R}$ is a 0-homogenous function.

Proof. Note that $\forall \epsilon > 0$, $\frac{\xi_\alpha}{|\xi|} \in C^\infty\big(\mathbb{R}^n \setminus B_\epsilon(0)\big)$, whence multiindex partial derivatives are well defined. Let $\boldsymbol{\beta}^\alpha := (\beta_1, \ldots, \beta_{\alpha-1}, \beta_\alpha + 1, \beta_{\alpha+1}, \ldots, \beta_n)$. We will use induction over $|\boldsymbol{\beta}|$ to proof that for all $\xi \in \mathbb{R}^n \setminus \{0\}$ the derivatives are a sum of terms of the form

$$c_{\alpha,\boldsymbol{\gamma},\boldsymbol{\beta}} \frac{\xi^{\boldsymbol{\gamma}}}{|\xi|^{|\boldsymbol{\gamma}|+|\boldsymbol{\beta}|}}, \tag{2.7}$$

where $\boldsymbol{\gamma} \le \boldsymbol{\beta}^\alpha$ and $c_{\alpha,\boldsymbol{\gamma},\boldsymbol{\beta}} \in \mathbb{C}$.

This is clear for $|\boldsymbol{\beta}| = 0$, since then $\boldsymbol{\beta}^\alpha = (\delta_{\alpha,k})_{k=1}^n$ which yields $\frac{\xi_\alpha}{|\xi|}$.

Let $|\boldsymbol{\beta}| = k$, $\nu \in \{1, \ldots, n\}$ and assume that $\partial^{\boldsymbol{\beta}} \frac{\xi_\alpha}{|\xi|}$ consists of terms of the form (2.7).

Then $|\boldsymbol{\beta}^\nu| = k+1$ and $\partial^{\boldsymbol{\beta}^\nu} \frac{\xi_\alpha}{|\xi|} = \frac{\partial}{\partial \xi_\nu} \partial^{\boldsymbol{\beta}} \frac{\xi_\alpha}{|\xi|}$. Hence we only need to proof that any partial derivative of degree one of a term of the form (2.7) is again a sum of terms of this form.

Let $\boldsymbol{\gamma} \le \boldsymbol{\beta}^\alpha$. By the product rule

$$\frac{\partial}{\partial \xi_\nu} \frac{\xi^{\boldsymbol{\gamma}}}{|\xi|^{|\boldsymbol{\gamma}|+|\boldsymbol{\beta}|}} = \frac{\xi^{(\gamma_1,\ldots,\gamma_{\nu-1},\gamma_\nu+1,\gamma_{\nu+1},\ldots,\gamma_n)}}{|\xi|^{|\boldsymbol{\gamma}|+2+|\boldsymbol{\beta}|}} + (1-\delta_{\gamma_\nu,0})\gamma_\nu \frac{\xi^{(\gamma_1,\ldots,\gamma_{\nu-1},\gamma_\nu-1,\gamma_{\nu+1},\ldots,\gamma_n)}}{|\xi|^{|\boldsymbol{\gamma}|+|\boldsymbol{\beta}|}}$$

Notice that for any multiindex $\boldsymbol{\mu} \in \mathbb{N}_0^n$ the function $\mathbb{R}^n \to \mathbb{R}$, $\xi \mapsto \frac{\xi^{\boldsymbol{\mu}}}{|\xi|^{|\boldsymbol{\mu}|}}$ is 0-homogenous. Now in the first term the degree of the monomial is increased by one, whereas the power in the denominator is increased by two. In the second term, which only exists if $\gamma_\nu \neq 0$, the degree of the monomial in the numerator is decreased by one, whereas the power in the denominator remains unchanged. As a consequence both terms are of the form $\frac{\xi^{\boldsymbol{\mu}}}{|\xi|^{|\boldsymbol{\mu}|+|\boldsymbol{\beta}^\nu|}}$, where $\boldsymbol{\mu} \le \boldsymbol{\beta}^\nu$. □

2.2.3 Riesz transforms of distributions modulo polynomials

Theorem 2.2.9 (Riesz transform and the Lizorkin space)
LET $\alpha \in \{1, \ldots, n\}$.
THEN *the partial Riesz transform $R_\alpha : \Phi(\mathbb{R}^n) \to \Phi(\mathbb{R}^n)$ is a bounded linear mapping.*

Proof. Let $\phi \in \Phi(\mathbb{R}^n)$, which is equivalent to

$$\widehat{\phi} \in \Psi(\mathbb{R}^n) = \big\{\psi \in \mathscr{S}(\mathbb{R}^n) : \partial^\gamma \psi(0) = 0 \ \forall \gamma = \{\gamma_1, \ldots, \gamma_n\} \in \mathbb{N}_0^n\big\}.$$

Remember that the partial Riesz transform R_α corresponds to a multiplication by the Fourier multiplier $\widehat{R_\alpha} = \frac{i\xi_\alpha}{|\xi|}$ in the Fourier domain.
To show that

$$R_\alpha \phi \in \Phi(\mathbb{R}^n)$$

we will show that

$$\widehat{R_\alpha \phi} \in \Psi(\mathbb{R}^n).$$

We start by proving $\widehat{R_\alpha\phi} \in C^\infty(\mathbb{R}^n)$, $\forall\phi \in \Phi(\mathbb{R}^n)$. Since $\forall\epsilon > 0$ the Fourier multiplier $\widehat{R_\alpha}$ is an element of $C^\infty(\mathbb{R}^n \setminus \overline{B_\epsilon(0)})$ it is clear that $\widehat{R_\alpha}\widehat{\phi} \in C^\infty(\mathbb{R}^n \setminus \overline{B_\epsilon(0)})$, $\forall\epsilon > 0$.

Furthermore $\widehat{R_\alpha} = \frac{i\xi_\alpha}{|\xi|} \in C(\mathbb{R}^n \setminus \{0\})$, whence $\widehat{R_\alpha\phi} \in C(\mathbb{R}^n \setminus \{0\})$.

For $\xi = 0$ we set $\widehat{R_\alpha\phi}(0) = 0$. Since the Fourier multiplier of the Riesz transform is $\frac{i\xi_\alpha}{|\xi|}$ and since $\xi_\alpha\phi(\xi) \in \Psi(\mathbb{R}^n)$ by Lemma 2.2.5 it follows that $\lim_{\xi\to 0}\widehat{R_\alpha\phi}(\xi) = \lim_{\xi\to 0}\frac{\xi_\alpha\widehat{\phi}(\xi)}{|\xi|} = 0$. Hence $\widehat{R_\alpha\phi} \in C(\mathbb{R}^n)$.

Let $\boldsymbol{\gamma} = (\gamma_1,\dots,\gamma_n) \in \mathbb{N}_0^n$, using Remark 2.2.7 and letting p, q as in Lemma 2.2.5

$$\begin{aligned}
\lim_{\xi\to 0}\widehat{R_\alpha\phi}^{(\boldsymbol{\gamma})}(\xi) &= \lim_{\xi\to 0}\sum_{k_1=0}^{\gamma_1}\cdots\sum_{k_n=0}^{\gamma_n}\prod_{l=1}^{n}\binom{\gamma_l}{k_l}\widehat{R_\alpha}^{(\boldsymbol{\gamma}-k)}(\xi)\phi^{(k)}(\xi)\\
&= \lim_{\xi\to 0}\sum_{k_1=0}^{\gamma_1}\cdots\sum_{k_n=0}^{\gamma_n}\prod_{l=1}^{n}\binom{\gamma_l}{k_l}\frac{p_{\alpha,\boldsymbol{\gamma}-k}(\xi)\widehat{\phi}^{(k)}(\xi)}{q_{2|\boldsymbol{\gamma}-k|+1}(\xi)}\\
&= 0,
\end{aligned}$$

where $k = (k_1,\dots,k_n) \in \mathbb{N}_0^n$.

To show continuity of $\widehat{R_\alpha\phi}^{(\boldsymbol{\gamma})}$ on $\mathbb{R}^n$ it remains to show $\widehat{R_\alpha\phi}^{(\boldsymbol{\gamma})}(0) = 0$, $\forall\boldsymbol{\gamma} \in \mathbb{N}_0^n$. This can be done inductively. We have already shown this for $\boldsymbol{\gamma} = 0$. Assume $\widehat{R_\alpha\phi}^{(\boldsymbol{\gamma})}(0) = 0$ has already been proven and let $\boldsymbol{\gamma}^{l+}$ be the multiindex such that $\boldsymbol{\gamma}_j^{l+} := \begin{cases}\gamma_j + 1 & \text{if } j = l\\ \gamma_j & \text{if } j \neq l.\end{cases}$.

Then

$$\widehat{R_\alpha\phi}^{(\boldsymbol{\gamma}^{l+})}(0) = \lim_{h\to 0}\left(\widehat{R_\alpha\phi}^{(\boldsymbol{\gamma})}(0) - \widehat{R_\alpha\phi}^{(\boldsymbol{\gamma})}(he_l)\right)/h = 0.$$

Finally to proof $\widehat{R_\alpha\phi} \in C^\infty(\mathbb{R}^n)$ we show that $\widehat{R_\alpha\phi}^{(\boldsymbol{\gamma})}(\xi)$ is bounded on $\mathbb{R}^n$. Let $\epsilon > 0$. Then $\widehat{R_\alpha\phi}^{(\boldsymbol{\gamma})}(\xi)$ is clearly bounded for $\xi \in \mathbb{R}^n\setminus(\overline{B_\epsilon(0)})$. Since we have shown that $\widehat{R_\alpha\phi}^{(\boldsymbol{\gamma})}$ is continuous on the compact set $\overline{B_\epsilon(0)}$ we know that it is bounded on this set, too.

Thus we have shown $\widehat{R_\alpha\phi} \in C^\infty(\mathbb{R}^n)$.

Next we will show that $R_\alpha\phi \in \mathscr{S}(\mathbb{R}^n)$. Let $\boldsymbol{\gamma} \in \mathbb{N}_0^n$ and $\beta \in \mathbb{N}_0$.

$$\begin{aligned}
\sup_{\xi\in\mathbb{R}^n}(1+|\xi|)^\beta|\partial^{\boldsymbol{\gamma}}\widehat{R_\alpha(\phi)}(\xi)| &= \sup_{\xi\in\mathbb{R}^n}(1+|\xi|)^\beta\left|\partial^{\boldsymbol{\gamma}}\left(i\frac{\xi_\alpha}{|\xi|}\widehat{\phi}\right)(\xi)\right|\\
&= \sup_{\xi\in\mathbb{R}^n}(1+|\xi|)^\beta\left|\sum_{k_1=0}^{\gamma_1}\cdots\sum_{k_n=0}^{\gamma_n}\prod_{j=1}^{n}\binom{\gamma_j}{k_j}\partial^{\boldsymbol{\gamma}-k}\frac{\xi_\alpha}{|\xi|}\partial_j^k\widehat{\phi}(\xi)\right|\\
&= \sup_{\xi\in\mathbb{R}^n}(1+|\xi|)^\beta\left|\sum_{k_1=0}^{\gamma_1}\cdots\sum_{k_n=0}^{\gamma_n}\prod_{j=1}^{n}\binom{\gamma_j}{k_j}p_{\alpha,\boldsymbol{\gamma}-k}(\xi)\frac{\partial_j^{k_j}\widehat{\phi}(\xi)}{q_{2|\gamma|+1}(|\xi|)}\right|,
\end{aligned}\tag{2.8}$$

where $p_{\alpha,\boldsymbol{\gamma}}$ and $q_{2|\boldsymbol{\gamma}-k|+1}$ as in Remark 2.2.7.

For $|\xi| > 1$ (2.8) is clearly bounded since $\widehat{\phi} \in \mathscr{S}(\mathbb{R}^n)$ and $\frac{p_{\alpha,\gamma}}{q_{2|\beta|+1}}$ is bounded. Now (2.8) is continuous on the compact set $\{\xi \in \mathbb{R}^n : |\xi| \leq 1\}$ and hence it is bounded.

This shows that $\widehat{R_\alpha\phi} \in \Psi(\mathbb{R}^n)$ and thus that

$$R_\alpha\phi \in \Phi(\mathbb{R}^n),\ \forall\phi \in \Phi(\mathbb{R}^n).$$

□

Definition 2.2.10 (Riesz transform of a distribution in Φ')
LET $f \in \Phi'(\mathbb{R}^n)$, $\alpha \in \{1,\dots,n\}$.
THEN *the* **Riesz transform** $R_\alpha : \Phi'(\mathbb{R}^n) \to \Phi'(\mathbb{R}^n)$ *is defined by*

$$R_\alpha f(\phi) := f(-R_\alpha \phi). \tag{2.9}$$

Corollary 2.2.11 (The Riesz transform of a distribution in Φ' is well defined)
LET $\alpha \in \{1,\dots,n\}$.
THEN *the Riesz transform* $R_\alpha : \Phi'(\mathbb{R}^n) \to \Phi'(\mathbb{R}^n)$ *as defined in Definition 2.2.10 is well defined.*

Proof. We will first show that the definition is coherent with our definition of the Riesz transform on $L^p(\mathbb{R}^n)$, $1 < p < \infty$. Let $1 < p < \infty$, $f \in L^p(\mathbb{R}^n)$, and let $\phi \in \Phi(\mathbb{R}^n)$.

$$\begin{aligned}
R_\alpha f(\phi) &= \int_{\mathbb{R}^n} R_\alpha f(x)\phi(x)\,dx \\
&= \int_{\mathbb{R}^n} i\frac{x_\alpha}{|x|}\widehat{f}(x)\mathscr{F}^{-1}(\phi)(x)\,dx \\
&= \int_{\mathbb{R}^n} \widehat{f}(x)\, i\frac{x_\alpha}{|x|}\widehat{\phi}(-x)\,dx \\
&= \int_{\mathbb{R}^n} \widehat{f}(-x)\frac{-ix_\alpha}{|x|}\widehat{\phi}(x)\,dx \\
&= \int_{\mathbb{R}^n} \mathscr{F}^{-1}(f)(x)\widehat{-R_\alpha\phi}(x)\,dx \\
&= f(-R_\alpha\phi).
\end{aligned}$$

Secondly we need to show that $R_\alpha(\Phi(\mathbb{R}^n)) \subseteq \Phi(\mathbb{R}^n)$. This was shown in Theorem 2.2.9 and hence $\forall \phi \in \Phi:\ R_\alpha\phi \in \Phi(\mathbb{R}^n)$.
Finally we have to proof that $R_\alpha f$ is continuous. Let $\{\phi_k\}_{k\in\mathbb{N}} \subset \Phi(\mathbb{R}^n)$ such that $\lim_{k\to\infty}\phi_k = 0$. In order to show that $R_\alpha f$ is continuous we have to show that $\lim_{k\to\infty} R_\alpha f(\phi_k) = 0$. But since R_α is linear and bounded and hence continuous on $\Phi(\mathbb{R}^n)$ it follows that $\lim_{k\to\infty} R_\alpha\phi_k = 0$ and hence $\lim_{k\to\infty} R_\alpha f(\phi_k) = -\lim_{k\to\infty} f(R_\alpha\phi) = 0$ since f is continuous.
Altogether it follows that $Rf \in \Phi'$. □

2.2.4 Riesz transforms of certain tempered distributions

While the Riesz transform is not well defined for all tempered distributions it is well defined for those whose support does not contain 0.

Definition 2.2.12 (Riesz transform of tempered distributions)
LET $f \in \mathscr{S}'(\mathbb{R}^n)$ *such that* $0 \notin \operatorname{supp}\widehat{f}$.
Let $\tau \in C^\infty(\mathbb{R}^n)$ *such that*

$$0 \notin \operatorname{supp}(\tau),\ \tau(x) = 1,\ \forall x \in \operatorname{supp}(\widehat{f}). \tag{2.10}$$

THEN *the* **Riesz transform**

$$R_\alpha : \{f \in \mathscr{S}'(\mathbb{R}^n) : 0 \notin \operatorname{supp}(f)\} \to \{f \in \mathscr{S}'(\mathbb{R}^n) : 0 \notin \operatorname{supp}(f)\}$$

is defined by

$$R_\alpha f(\phi) = \widehat{f}\big(\tau\mathscr{F}^{-1}(-R_\alpha\phi)\big),\ \forall \phi \in \{f \in \mathscr{S}(\mathbb{R}^n) : 0 \notin \operatorname{supp}(f)\}. \tag{2.11}$$

Theorem 2.2.13 (Riesz transform of a distribution in $\mathscr{S}'$)
LET $f \in \mathscr{S}'(\mathbb{R}^n)$ *such that* $0 \notin \operatorname{supp}(\widehat{f})$ *and* $\tau \in C^\infty(\mathbb{R}^n)$ *such that* (2.10) *holds.*
THEN *the Riesz transform defined by*

$$R_\alpha f(\phi) = \widehat{f}\big(\tau \mathscr{F}^{-1}(-R_\alpha \phi)\big), \; \forall \phi \in \mathscr{S}(\mathbb{R}^n) \tag{2.12}$$

is well defined and independend on the choice of τ.

Proof. Since the support of f is a closed set that does not contain 0 there is a closed set containing 0 in the complement of $\operatorname{supp}(f)$. Hence it is indeed possible to choose τ as stated above.
We will proof that the Riesz transform is independent on the choice of τ. Let $\sigma \in C^\infty(\mathbb{R}^n)$ be another function such that $\sigma(x) = 1, \; \forall x \in \operatorname{supp}(\widehat{f}), \; \{0\} \cap \operatorname{supp}(\sigma) = \emptyset$. Furthermore let $\phi \in \mathscr{S}(\mathbb{R}^n)$. Then

$$\begin{aligned} &\widehat{f}\big(\tau \mathscr{F}^{-1}(-R_\alpha \phi)\big) - \widehat{f}\big(\sigma \mathscr{F}^{-1}(-R_\alpha \phi)\big) \\ &= \widehat{f}\big((\tau - \sigma)\mathscr{F}^{-1}(-R_\alpha \phi)\big) \\ &\overset{*}{=} 0. \end{aligned}$$

To see that $*$ holds note that $(\tau - \sigma)\mathscr{F}^{-1}(\phi) \in \Psi(\mathbb{R}^n) \subset \mathscr{S}(\mathbb{R}^n)$ and thus by Theorem 2.2.9 that $(\tau - \sigma)\mathscr{F}^{-1}(-R_\alpha \phi) \in \Psi(\mathbb{R}^n) \subset \mathscr{S}(\mathbb{R}^n)$. $*$ now follows since $\operatorname{supp}((\tau - \sigma)(\mathscr{F}^{-1}(\phi)) \cap \operatorname{supp}(f) = \emptyset$.
The proof that this Riesz transformation is coherent with the Riesz transform on $L^p(\mathbb{R}^n)$, $1 < p < \infty$, is analogous to the proof of this fact in Definition 2.2.10.

□

2.2.5 Hilbert and Riesz transforms

The next theorem states a connection between the Hilbert transform of distributions in $\mathscr{S}'(\mathbb{R})$ and the Riesz transforms of certain distributions in $\mathscr{S}'(\mathbb{R}^n)$ which are derived from distributions in $\mathscr{S}'(\mathbb{R})$ using the operator $\mathscr{L}_d$ defined below. This operator maps a one dimensional regular distribution $g \in \mathscr{S}'(\mathbb{R})$ to a n dimensional regular distribution $\mathscr{L}_d(g) = g(\langle \cdot , d\rangle) \in \mathscr{S}'(\mathbb{R}^n)$. (See Example 2.2.1) $\mathscr{L}_d(g)$ is sometimes called a plane wave.

Proposition 2.2.14 (The operator $\mathscr{L}_d$)
LET $f, g \in \mathscr{S}'(\mathbb{R})$, $d \in S^{n-1}$ *and let* $\rho_d \in SO(n) : \rho_d(d) = e_n$.
Let $\mathscr{L}_d : \mathscr{S}'(\mathbb{R}) \to \mathscr{S}'(\mathbb{R}^n)$ *be defined by*

$$\mathscr{L}_d f(\phi) = \int_{\mathbb{R}^{n-1}} f(\rho(\phi)(x, \cdot))\, dx, \; \forall \phi \in \mathscr{S}(\mathbb{R}^n).$$

THEN $\widehat{\mathscr{L}_d(f)}(\phi) = \widehat{f}(\rho\phi(0_{\mathbb{R}^{n-1}}, \cdot))$.
Furthermore

$$\mathscr{L}_d(f) + \mathscr{L}_d(g) = \mathscr{L}_d(f + g).$$

If additionally $\operatorname{supp}(\widehat{g})$ *is bounded*

$$\mathscr{L}_d(f)\mathscr{L}_d(g) = \mathscr{L}_d(fg).$$

If $\operatorname{supp}(\widehat{f}) \subseteq \mathbb{R} \setminus B_\epsilon(0)$ *for some* $\epsilon < 0$ *then*

$$\operatorname{supp}(\widehat{\mathscr{L}_d f}) \subseteq \left\{x \in \mathbb{R}^n : |\langle x, d\rangle| \geq \epsilon\right\}. \tag{2.13}$$

If $\operatorname{supp}(\widehat{f}) \subseteq B_\epsilon(0)$ *for some* $\epsilon > 0$ *then*

$$\operatorname{supp}(\widehat{\mathscr{L}_d f}) \subseteq \left\{x \in \mathbb{R}^n : |\langle x, d\rangle| \leq \epsilon\right\}.$$

Proof. We have to show that $\mathscr{L}_d$ is well defined, that is independent on the choice of ρ_d and that it maps into $\mathscr{S}'(\mathbb{R}^n)$.

To show the last it suffices to show that $\mathscr{L}(f)$ is a direct product of distributions. Then Theorem A.2.7 states the result. Let us define for a rotation $\rho \in SO(n)$ and a distribution $g \in \mathscr{S}'(\mathbb{R}^n)$ the operator $\rho : \mathscr{S}'(\mathbb{R}^n) \to \mathscr{S}'(\mathbb{R}^n)$ via $\rho(g)(\phi) := g(\rho\phi)$. The operator ρ is obviously well defined, since the operator $\rho : \mathscr{S}(\mathbb{R}^n) \to \mathscr{S}(\mathbb{R}^n)$ introduced in the preliminaries in Definition 1.3.21 maps test functions to test functions.

Let $I(\phi) := \int_{\mathbb{R}^{n-1}} \phi(x)dx,\ \forall \phi \in \mathscr{S}(\mathbb{R}^{n-1})$. Then $\mathscr{L}_d = \rho_d(I \times f)$ and hence well defined as a direct product of distributions.

We will now show that our definition is independent of the choice of ρ_d. Let $\rho, \sigma \in SO(n)$ be two rotations such that $\rho(d) = \sigma(d) = e_n$. Than there exists a rotation $\tau \in SO(n-1)$, such that $\rho = \sigma \begin{pmatrix} \tau & 0 \\ 0 & 1 \end{pmatrix}$. Hence

$$\begin{aligned}
\int_{\mathbb{R}^{n-1}} f(\sigma(\phi)(x,\ \cdot\ ))dx &= \int_{\mathbb{R}^{n-1}} f(\sigma\phi(\tau y,\ \cdot\ ))dy \\
&= \int_{\mathbb{R}^{n-1}} f(\sigma \begin{pmatrix} \tau & 0 \\ 0 & 1 \end{pmatrix} \phi(y,\ \cdot\ ))dy \\
&= \int_{\mathbb{R}^{n-1}} f(\rho\phi(y,\ \cdot\ ))dy \\
&= \mathscr{L}_d(f)(\phi),\ \forall \phi \in \mathscr{S}(\mathbb{R}^n).
\end{aligned}$$

Using Theorem 1.3.12 and Theorem 1.3.22

$$\begin{aligned}
\widehat{\mathscr{L}_d(f)}(\phi) &= \rho(I \times f)(\widehat{\phi}) = I \times f(\rho\widehat{\phi}) = \widehat{I} \times \widehat{f}(\rho\phi) \\
&= \widehat{f}(\rho\phi(0_{\mathbb{R}^{n-1}},\ \cdot\ )).
\end{aligned}$$

Let $g \in \mathscr{S}'(\mathbb{R}^n)$. Then

$$\begin{aligned}
\Big(\mathscr{L}_d(f) + \mathscr{L}_d(g)\Big)(\phi) &= \int_{\mathbb{R}^{n-1}} f(\rho_d\phi(x,\ \cdot\ ))dx + \int_{\mathbb{R}^{n-1}} g(\rho_d\phi(x,\ \cdot\ ))dx \\
&= \int_{\mathbb{R}^{n-1}} (f+g)(\rho_d\phi(x,\ \cdot\ ))dx \\
&= \mathscr{L}_d(f+g)(\phi).
\end{aligned}$$

Let furthermore $\operatorname{supp}(\widehat{g})$ be bounded. By Corollary A.2.13 fg is well defined, by Theorem A.2.11 $g \in C^\infty(\mathbb{R}^n)$ and by Theorem A.2.10 $fg(\phi) = f(g\phi),\ \forall \phi \in \mathscr{S}(\mathbb{R}^n)$. Hence

$$\begin{aligned}
\Big(\mathscr{L}_d(g)\mathscr{L}_d(f)\Big)(\phi) &= \int_{\mathbb{R}^{n-1}} f\Big(\rho_d\big(g(\langle(x,\ \cdot\ ), d\rangle)\phi(x,\ \cdot\ )\big)\Big)dx \\
&= \int_{\mathbb{R}^{n-1}} f\Big(g(\ \cdot\ )\rho_d\big(\phi(x,\ \cdot\ )\big)\Big)dx \\
&= \mathscr{L}_d(gf)(\phi),\ \forall \phi \in \mathscr{S}(\mathbb{R}^n).
\end{aligned}$$

To proof (2.13) note that

$$x \in \operatorname{supp}(\widehat{\mathscr{L}_d(f)}) \Leftrightarrow \left(\rho^{-1}x\right)_n \in B_\epsilon(0) \Leftrightarrow |\langle x, d\rangle| \in B_\epsilon(0)$$

□

The following example will illustrate the effect of the operator $\mathscr{L}_d$.

EXAMPLE 2.2.1:
Let $f \in C^\infty(\mathbb{R})$. Then f defines a distribution in $\mathscr{S}'(\mathbb{R})$ via

$$f(\phi) = \int_{\mathbb{R}} f(t)\phi(t)dt,\ \forall \phi \in \mathscr{S}(\mathbb{R}).$$

It follows that

$$\mathscr{L}_{e_n}(f)(\phi) = \int_{\mathbb{R}^{n-1}} \int_{\mathbb{R}} f(x_n)\phi(x_1, \dots, x_{n-1}, x_n)dx_1 \dots dx_{n-1}dx_n,\ \forall \phi \in \mathscr{S}(\mathbb{R}^n).$$

And

$$\begin{aligned}\mathscr{L}_d(f)(\phi) &= \int_{\mathbb{R}^{n-1}} \int_{\mathbb{R}} f(x_n)\rho_d\phi(x_1, \dots, x_{n-1}, x_n)dx_1 \dots dx_{n-1}dx_n = \int_{\mathbb{R}^n} f((\rho^{-1}y)_n)\phi(y)dy \\ &= \int_{\mathbb{R}^n} f(\langle y, d\rangle)\phi(y)dy,\ \forall \phi \in \mathscr{S}(\mathbb{R}^n).\end{aligned}$$

That is, in distributional sense $\mathscr{L}_d(f) = f(\langle\ \cdot\ , d\rangle)$.

Theorem 2.2.15 (Hilbert and Riesz transforms)
LET $f \in \mathscr{S}'(\mathbb{R})$ *such that* $0 \notin \operatorname{supp}(\widehat{f})$. *Furthermore let* $d \in \mathbb{R}^n : \|d\| = 1$.
THEN *the Riesz transform of* $\mathscr{L}_d(f) \in \mathscr{S}'(\mathbb{R}^n)$ *exists. Indeed*

$$R_d\mathscr{L}_d(f) = \mathscr{L}_d(\mathscr{H}f),$$

and

$$R_l\mathscr{L}_d(f) = 0,\ \forall l \in \mathbb{R}^n : \langle d, l\rangle = 0.$$

Proof. Since $\mathbb{R} \setminus \{0\}$ is an open set, while $\operatorname{supp}(\widehat{f})$ is closed, $\exists \epsilon > 0 : B_\epsilon(0) \cap \operatorname{supp}(\widehat{f}) = \emptyset$. It follows that $0 \in \{x \in \mathbb{R}^n : |\langle x, d\rangle| < \epsilon\}$ and by (2.13) it follows that

$$\left\{x \in \mathbb{R}^n : |\langle x, d\rangle| < \epsilon\right\} \cap \operatorname{supp}(\widehat{\mathscr{L}_d(f)}) = \emptyset.$$

Let us choose $\tau \in C^\infty(\mathbb{R})$ satisfying (2.10) for $f \in \mathscr{S}'(\mathbb{R})$, that is $\operatorname{supp}(\tau) \cap \{0\} = \emptyset$ and $\tau(x) = 1,\ \forall x \notin B_\epsilon(0)$. Then $\tau_d(x) = \tau(\langle x, d\rangle)$ satisfies (2.10) and we can apply Definition 2.2.12 to $\mathscr{L}_d(f)$.

$$\begin{aligned}R_d\mathscr{L}_d(f)(\phi) &= \widehat{\mathscr{L}_d(f)}\left(\widehat{-R_d}\tau_d\mathscr{F}^{-1}(\phi)\right) = \widehat{f}\left(\rho\left(-\widehat{R_d}\tau_d\mathscr{F}^{-1}(\phi)\right)(0_{\mathbb{R}^{n-1}},\ \cdot\ )\right) \\ &= \widehat{f}\left(i\frac{-\ \cdot}{|(0_{\mathbb{R}^{n-1}},\ \cdot\ )|}\tau\left(\langle(0_{\mathbb{R}^{n-1}},\ \cdot\ ), e_n\rangle\right)\rho(\mathscr{F}^{-1}(\phi))(0_{\mathbb{R}^{n-1}},\ \cdot\ )\right) \\ &= \widehat{f}\left(i\operatorname{sgn}(\ \cdot\ )\tau(\ \cdot\ )\rho\left(\widehat{\phi}(0_{\mathbb{R}^{n-1}}, -\ \cdot\ )\right)\right) \\ &= \mathscr{L}_d(\mathscr{H}f)(\phi),\ \forall \phi \in \mathscr{S}(\mathbb{R}^n).\end{aligned}$$

Furthermore, for any $l \in S^{n-1}$ such that $\langle l, d\rangle = 0$ there exists some $a_l \in \mathbb{R}^{n-1}$ with $\rho_d(l) = (a_l, 0)$. Hence

$$\begin{aligned} R_l \mathscr{L}_d(f)(\phi) &= \widehat{\mathscr{L}_d(f)}\Big(-\widehat{R_l}\tau_d \mathscr{F}^{-1}(\phi)\Big) = \widehat{f}\Big(\rho\big(-\widehat{R_l}\tau_d\phi\big)(0_{\mathbb{R}^{n-1}}, \cdot\,)\Big) \\ &= \widehat{f}\Big(i\tfrac{0_{\mathbb{R}}}{|(0_{\mathbb{R}^{n-1}}, \cdot\,)|}\tau\big(\langle(0_{\mathbb{R}^{n-1}}, \cdot\,), e_n\rangle\big)\rho(\phi)(0_{\mathbb{R}^{n-1}}, \cdot\,)\Big) \\ &= \widehat{f}\Big(0_{C^\infty(\mathbb{R}^n)}\rho(\phi)(0_{\mathbb{R}^{n-1}}, \cdot\,)\Big) \\ &= 0_{\mathscr{S}'(\mathbb{R}^n)}(\phi), \ \forall \phi \in \mathscr{S}(\mathbb{R}^n). \end{aligned}$$

□

2.2.6 The Bedrosian identity and Riesz transforms

The Bedrosian identity has first been proven by Bedrosian in [5]. The purpose is to determine under which conditions the amplitude phase decomposition achieved by the analytical signal gives the phase and amplitude one would expect by the construction of the signal. That is, let $a \in L^2(\mathbb{R})$, $\phi : \mathbb{R} \to \mathbb{R} :\ \cos(\phi) \in L^2(\mathbb{R})$ and let $f = a\cos(\phi) \in L^2(\mathbb{R})$. The Bedrosian identity states conditions on a and $\cos(\phi)$ under which $f_a = a\exp(\phi)$.

In this section we will extend the Bedrosian identities to distributions and give a Bedrosian identity for Riesz transforms.

Let us however first state the classical Bedrosian identity [5].

Theorem 2.2.16 (The Bedrosian identity)
LET *$f, g \in L^2(\mathbb{R})$ such that* $\operatorname{supp}(\widehat{f}) \subseteq [a, b]$ *and* $\operatorname{supp}(\widehat{g}) \subset \mathbb{R} \setminus [-b, -a]$.
THEN

$$\mathscr{H}(fg) = f\mathscr{H}(g). \tag{2.14}$$

Remark 2.2.17
The original Bedrosian identity is given for functions in $f, g \in L^2(\mathbb{R})$. It is at first view not clear if $\mathscr{H}fg \in L^2(\mathbb{R})$ or in which sense this identity should be interpreted, it is simply assumed to be well defined in some sense.

Indeed since $\operatorname{supp}(\widehat{f})$ *is bounded, $\widehat{f} \in L^1(\mathbb{R})$, and by Corollary A.2.12 it holds that $f \in C^\infty(\mathbb{R})$ is bounded and thus $fg \in L^2(\mathbb{R}^n)$.*

In the case of a Bedrosian identity for distributions we have to be more careful: It does hold true that $fg \in \mathscr{S}'(\mathbb{R})$ for $f \in \mathscr{S}'(\mathbb{R}^n)$ with compact support in the Fourier domain and $g \in \mathscr{S}'(\mathbb{R})$. But $fg \in \mathscr{S}'$ is not enough to ensure that the Riesz transform is well defined. It is remarkable that this well definedness is guaranteed by the same condition on the support of the distributions that is needed for the Bedrosian identity to hold.

Theorem 2.2.18 (The Bedrosian identity for distributions)
LET *$a < 0 < b$ and let $f \in \mathscr{S}'(\mathbb{R})$ such that* $\operatorname{supp}(\widehat{f}) \subseteq [a, b]$ *and $g \in \mathscr{S}'(\mathbb{R})$ such that* $\operatorname{supp}(\widehat{g}) \subset \mathbb{R} \setminus [-b, -a]$.
THEN

$$\mathscr{H}(fg) = f\mathscr{H}(g). \tag{2.15}$$

Proof. Since the support of a distribution is a closed set,

$$\exists \epsilon > 0 : \operatorname{supp}(\widehat{g}) \subseteq \mathbb{R} \setminus]-b-\epsilon, -a+\epsilon[.$$

Let $\phi \in \mathscr{S}(\mathbb{R})$.

First we have to show that $\mathscr{H} fg$ and $f\mathscr{H}g$ are well defined in $\mathscr{S}'(\mathbb{R})$.

Since $\operatorname{supp}(\widehat{f})$ is bounded the convolution $\widehat{f} * \widehat{g}$ is well defined by Theorem A.2.14 as a direct product. Furthermore $0 \notin \operatorname{supp}(\widehat{fg})$:

$$\widehat{fg}(\phi) = \widehat{f} * \widehat{g}(\phi) = \widehat{f} \times \widehat{g}\big(\phi(\,\cdot\,_f + \,\cdot\,_g)\big),\ \forall \phi \in \mathscr{S}(\mathbb{R}^n)$$

If $x + y \in]-\epsilon, \epsilon[$ then either $x \notin [a,b] = \operatorname{supp}(\widehat{f})$ or $y \notin \mathbb{R} \setminus]-b-\epsilon, -a+\epsilon[= \operatorname{supp}(\widehat{g})$. Hence $0 \notin \operatorname{supp}(\widehat{fg})$.

Hence fg satisfies the condition of Theorem 2.2.13, which states that $\mathscr{H} fg \in \mathscr{S}'(\mathbb{R})$.

Now g satisfies the conditions of Theorem 2.2.13 as well, that is $\mathscr{H}g \in \mathscr{S}'(\mathbb{R})$ and $f\mathscr{H}g \in \mathscr{S}'(\mathbb{R})$ by Theorem A.2.14.

Now let $\sigma_f, \tau_g, \mu_{fg} \in C^\infty(\mathbb{R})$ such that

$$\begin{aligned}
&\sigma_f(x) = 1\ \forall x \in [a,b], && \operatorname{supp}(\sigma_f) \subseteq [a-\epsilon/2, b+\epsilon/2],\\
&\tau_g(y) = 1,\ \forall y \notin]-b-\epsilon, -a+\epsilon[, && \operatorname{supp}(\tau_g) \cap [-b-\epsilon/2, -a+\epsilon/2] = \emptyset,\\
&\mu_{fg}(z) = 1,\ \forall z \notin [-\epsilon, \epsilon], && \operatorname{supp}(\mu_{fg}) \cap [-\epsilon/2, \epsilon/2] = \emptyset.
\end{aligned}$$

It follows that $\kappa_x(y) := \mu_{fg}(x+y)\sigma_f(x) \in C^\infty(\mathbb{R})$ and $\kappa_x(y) = 1,\ \forall y \in [-\epsilon-b, -a+\epsilon]$. Consequently σ_f satisfies (2.10) for f, $\tau_g\kappa_x$ satisfies (2.10) for g and μ_{fg} satisfies (2.10) for fg.

$$\begin{aligned}
\mathscr{H} fg(\phi) &\overset{2.2.12}{=} \widehat{fg}\big(-\widehat{\mathscr{H}}\mu_{fg}\widehat{\phi}^\vee\big) \overset{A.2.14}{=} \widehat{f} * \widehat{g}\big(-\widehat{\mathscr{H}}\mu_{fg}\widehat{\phi}^\vee\big) &(2.16)\\
&\overset{(1)}{=} \widehat{f}\Big(\widehat{g}\big(-i\operatorname{sgn}(\,\cdot\,_f + \,\cdot\,_g)\mu_{fg}(\,\cdot\,_f + \,\cdot\,_g)\widehat{\phi}(-\,\cdot\,_f - \,\cdot\,_g)\big)\Big)\\
&\overset{(2)}{=} \widehat{f}\Big(\widehat{g}\big(-i\operatorname{sgn}(\,\cdot\,_f + \,\cdot\,_g)\sigma_f(\,\cdot\,_f)\tau_g(\,\cdot\,_g)\mu_{fg}(\,\cdot\,_f + \,\cdot\,_g)\widehat{\phi}(-\,\cdot\,_f - \,\cdot\,_g)\big)\Big)\\
&\overset{(3)}{=} \widehat{f}\Big(\widehat{g}\big(-i\operatorname{sgn}(\,\cdot\,_g)\sigma_f(\,\cdot\,_f)\mu_{fg}(\,\cdot\,_f + \,\cdot\,_g)\tau_g(\,\cdot\,_g)\widehat{\phi}(-\,\cdot\,_f - \,\cdot\,_g)\big)\Big)\\
&= \widehat{f}\Big(\widehat{g}\big(-i\operatorname{sgn}(\,\cdot\,_g)\kappa_{\,\cdot\,_f}(\,\cdot\,_g)\tau_g(\,\cdot\,_g)\widehat{\phi}(-\,\cdot\,_f - \,\cdot\,_g)\big)\Big)\\
&\overset{2.2.12}{=} \widehat{f}\Big(\widehat{\mathscr{H}g}\big(\widehat{\phi}(-\,\cdot\,_f - \,\cdot\,_g)\big)\Big)\\
&\overset{(1)}{=} \widehat{f} * \widehat{\mathscr{H}g}(\widehat{\phi}^\vee)\\
&= f\mathscr{H}g(\phi),
\end{aligned}$$

where $\phi^\vee(x) = \phi(-x),\ \forall x \in \mathbb{R}$.

(1) is the convolution of the two distributions $\widehat{f}$ and $\widehat{\mathscr{H}g}$. This convolution is well defined since $\widehat{f}$ has bounded support. (See Theorem A.2.14.)

(2) We used the fact that if $h \in \mathscr{S}'(\mathbb{R})$, $\nu \in C^\infty(\mathbb{R}) : \nu(x) = 1,\ \forall x \in \operatorname{supp}(h)$ and $\phi \in \mathscr{S}'(\mathbb{R})$ then

$$h(\phi) = h(\nu\phi + (1-\nu)\phi) = h(\nu\phi).$$

(3) holds true since $\operatorname{sgn}(x+y) = \operatorname{sgn}(y)$, whenever

$$x + y > 0 \wedge y > 0$$

or

$$x + y < 0 \wedge y < 0.$$

Now this holds true if $x \in \operatorname{supp}(\sigma_f) \subseteq [a,b]$ and $y \in \operatorname{supp}(\tau_g) \subseteq \mathbb{R}\setminus]-b-\epsilon, a+\epsilon[$.

□

EXAMPLE 2.2.2:
The calculation (2.16) is made somewhat intransparent by the necessary use of the functions σ, τ, μ and κ. Therefore we will now calculate the same fact under the additional assumption that $\widehat{f}, \widehat{g} \in L^1_{\text{loc}}$. Then the integration bounds will take over the role of the functions σ, τ, μ and κ.

$$\begin{aligned}
\mathscr{H} fg(\phi) &\overset{2.2.12}{=} \widehat{fg}\big(-\widehat{\mathscr{H}\mu_{fg}}\widehat{\phi}^{\vee}\big) \overset{A.2.14}{=} \widehat{f} * \widehat{g}\big(-\widehat{\mathscr{H}\mu_{fg}}\widehat{\phi}^{\vee}\big) \\
&\overset{(1)}{=} \int_{[a,b]} \widehat{f}(x) \int_{\mathbb{R}\setminus[-\epsilon-a,\epsilon-b]} -i\operatorname{sgn}(x+y)\widehat{g}(y)\underbrace{\mu_{fg}(x+y)}_{=1}\widehat{\phi}(-x-y)\,dx\,dy \\
&\overset{(2)}{=} \int_{[a,b]} \widehat{f}(x) \int_{\mathbb{R}\setminus]-\epsilon-b,\epsilon-a[} -i\operatorname{sgn}(y)\widehat{g}(y)\widehat{\phi}(-x-y)\,dx\,dy \\
&\overset{2.2.12}{=} \widehat{f}\big(\widehat{\mathscr{H}g}\big(\widehat{\phi}(-\cdot_f - \cdot_g)\big)\big) \\
&\overset{(1)}{=} \widehat{f} * \widehat{\mathscr{H}g}(\widehat{\phi}^{\vee}) \\
&= f\mathscr{H}g(\phi)
\end{aligned}$$

(1) is the convolution of the two distributions $\widehat{f}$ and $\widehat{\mathscr{H}g}$. This convolution is well defined since $\widehat{f}$ has bounded support. (See Theorem A.2.14.)

(2) holds true since $\operatorname{sgn}(x+y) = \operatorname{sgn}(y)$, whenever $x+y > 0 \wedge y > 0$ or $x+y<0 \wedge y<0$. Now this holds true since $\operatorname{supp}(\widehat{f}) \subseteq [a,b]$ and $\operatorname{supp}(g) \subseteq \mathbb{R}\setminus]-b-\epsilon, a+\epsilon[$.

Theorem 2.2.19 (A Bedrosian identity for the Riesz transform)
LET $f, g \in \mathscr{S}'(\mathbb{R})$ *such that* $\operatorname{supp}(\widehat{f}) \subseteq [-R,R]$ *and* $\operatorname{supp}(\widehat{g}) \in \mathbb{R}\setminus[-R,R]$ *for some* $R \in \mathbb{R}_+$. *Furthermore let* $d \in S^{n-1}$.

THEN $F = \mathscr{L}_d f$ *and* $G = \mathscr{L}_d g$ *are in* $\mathscr{S}'(\mathbb{R}^n)$ *and satisfy the Bedrosian identity*

$$R(FG) = FR(G).$$

Proof. We will first show that $F \in C^\infty(\mathbb{R}^n)$. By Theorem A.2.11 $f \in C^\infty(\mathbb{R})$ since the support of its Fourier transform is bounded. Furthermore $|f(x)| \leq \gamma(1+|x|)^N$, $\forall x \in \mathbb{R}$, where N is the order of the distribution f which is finite. Since derivatives in the time respectively space domain correspond to multiplication by a polynomial on the Fourier domain Theorem 1.3.17, we can once again use Theorem A.2.11 to get

$$\forall p \in \mathbb{N}\, \exists N \in \mathbb{N}, \gamma_{p,N} > 0 : |D^p f(x)| \leq \gamma_{p,N}(1+|x|)^N,\ \forall x \in \mathbb{R}. \tag{2.17}$$

If $f \in C^\infty(\mathbb{R})$, then $F = f(\langle \cdot, d\rangle) \in C^\infty(\mathbb{R}^n)$. By (2.17) and since $d \in S^{n-1}$ it follows that

$$\forall \alpha \in \mathbb{N}_0^n\ |D^\alpha F(x)| \leq \gamma_{p,N}(1+|x|)^{N_p},\ \forall x \in \mathbb{R}^n,$$

where $p = |\alpha| = \sum_{l=1}^{n} \alpha_l$. Furthermore, by construction $G \in \mathscr{S}'(\mathbb{R}^n)$. Thus by Theorem A.2.10(i) it follows that $FG \in \mathscr{S}'(\mathbb{R}^n)$ is well defined.

To proof that the Riesz transform of FG is well defined in $\mathscr{S}'(\mathbb{R}^n)$ it is sufficient to show that $0 \notin \operatorname{supp}(\widehat{FG})$.

Since $FG \in \mathscr{S}'(\mathbb{R}^n)$ by Theorem A.2.14 we know that $\widehat{FG} = \widehat{F} * \widehat{G}$. That is, we can proceed as in the proof of Theorem 2.2.18. We recall that, since the support of a distribution is a closed set $\exists \epsilon > 0 : \operatorname{supp}(\widehat{g}) \subseteq \mathbb{R} \backslash]-R-\epsilon, R+\epsilon[$. It follows that

$$\begin{aligned} \widehat{FG}(\phi) = \widehat{F} * \widehat{G}(\phi) &= \widehat{F}\Big(\widehat{G}\big(\phi(\,\cdot\,_F + \,\cdot\,_G)\big)\Big) \\ &= \widehat{f}\Big(\widehat{g}\big(\phi(\langle \,\cdot\,_F + \,\cdot\,_G, d\rangle)\big)\Big). \end{aligned}$$

And hence

$$\xi \in \operatorname{supp}(\widehat{FG}) \Leftrightarrow \langle \xi, d\rangle = y + z,$$

where $y \in \operatorname{supp}(\widehat{f}) \subseteq [-R, R]$ and $z \in \operatorname{supp}(\widehat{g}) \subseteq \mathbb{R} \backslash]-R-\epsilon, R+\epsilon[$. As a consequence $\xi \in \operatorname{supp}(\widehat{FG})$ if $\langle \xi, d\rangle \in \mathbb{R} \setminus [-\epsilon, \epsilon]$. Hence $0 \in \{\xi \in \mathbb{R}^n : \langle \xi, d\rangle \le \epsilon\} \not\subseteq \operatorname{supp}(\widehat{FG})$. Thus Theorem 2.2.13 shows that $R(FG) \in \mathscr{S}'(\mathbb{R}^n)$.

We will now show that $FRG \in \mathscr{S}'(\mathbb{R}^n)$ is well defined. Theorem 2.2.13 shows that $RG \in \mathscr{S}'(\mathbb{R}^n)$ is well defined. But then Theorem A.2.14 shows that $FRG \in \mathscr{S}'(\mathbb{R}^n)$ is well defined.

Thus it remains only to show $FRG = R(FG)$. To show this it is sufficient to show that $FR_dG = R_d(FG)$ since by Theorem 2.2.15 $R_l g$, respectively $R_l FG$ vanish for all $l \in \mathbb{R}^n : \langle l, d\rangle = 0$. Indeed

$$\begin{aligned} R_dFG - FR_dG &= R_d\mathscr{L}_d(fg) - \mathscr{L}_d(f)R_d\mathscr{L}_d(g) \\ &= \mathscr{L}_d(\mathscr{H}fg) - \mathscr{L}_d(f\mathscr{H}g) \\ &= \mathscr{L}_d(\mathscr{H}fg - f\mathscr{H}g) \\ &= 0, \end{aligned}$$

where we used Theorem 2.2.18, Theorem 2.2.15 and Proposition 2.2.14.

□

Remark 2.2.20
In [54] it was proven that a translation invariant operator on $L^2(\mathbb{R}^n)$ satisfies a Bedrosian theorem iff it is a linear combination of compositions of partial Hilbert transforms. That is the Fourier multiplier of the operator is of the form

$$\sum_{\boldsymbol{\alpha}} a_{\boldsymbol{\alpha}} \prod_{k=1}^{n} \frac{x_k^{\alpha_k}}{|x_k^{\alpha_k}|},$$

where $a_{\boldsymbol{\alpha}} \in \mathbb{C}$ and $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}_0^n$ is a multiindex. As a consequence a general Bedrosian theorem for the Riesz transform can not exist.

EXAMPLE 2.2.3 (The Bedrosian identity and phase decomposition of signals):
The analytical signal yields a unique decomposition of a function $f = a\cos(\varphi)$, where a is the amplitude and φ the phase of the analytical signal.

1. Let us now consider a function of the form

$$f(t) = a(t)\cos(\varphi(t)),$$

with $a(t), \cos(\varphi(t)) \in \mathbb{R},\ \forall t \in \mathbb{R}$ and $\int_{\mathbb{R}} a(t) \cdot (t) dt,\ \int_{\mathbb{R}} \cos(\varphi(t)) \cdot (t) dt \in \mathscr{S}'(\mathbb{R})$. We want to know under which conditions its analytical signal will be

$$f_a(t) = f(t) + i\mathscr{H} f(t) = a(t) \exp(i\varphi(t)),$$

that is a is the amplitude and φ is the phase.

Theorem 2.2.18 tells us that this is the case, if $\exists -\infty < c < 0 < d < \infty$ such that $\mathrm{supp}(\widehat{a}) \subseteq [c,d]$ and $\mathrm{supp}(\widehat{\cos(\varphi)}) \subset \mathbb{R} \setminus [-d,-c]$.

A possible choice is $a(t) = \cos(rt),\ \varphi(t) = st,\ \forall t \in \mathbb{R}$, where $0 < r < s < \infty$.

2. Theorem 2.2.19 answers a similar question for the Riesz transform respectively the monogenical signal defined in chapter 4:

   Considering a function
   $$F(x) = A(x)\cos(\Phi(x))$$
   such that $A(x), \cos(\Phi(x)) \in \mathbb{R},\ \forall x \in \mathbb{R}^n$ and A, $\cos(\Phi)$ define a distribution in $\mathscr{S}'(\mathbb{R}^n)$.

   Let
   $$F_m(x) = (F(x), RF(x)) = A\Big(\cos(\phi), \frac{RF}{|RF|}\sin(\phi)\Big) \in \mathbb{R}^{n+1}$$
   define the monogenic signal, where
   $$A(x) = |F(x), RF(x)|_{\mathbb{R}^{n+1}}$$
   and
   $$\Phi(x) = \arg\Big(\frac{F(x) + i|RF(x)|}{|F(x), RF(x)|_{\mathbb{R}^{n-1}}}\Big).$$
   We want to know under which conditions A will be the amplitude and Φ the phase of the monogenic signal, i.e.
   $$\exists d(x) \in \mathbb{R}^n :\ F_m(x) = A\Big(\cos(\Phi(x)) + d(x)\sin(\Phi(x))\Big).$$

   Theorem 2.2.19 states that this is the case if $F := \mathscr{L}_d(f)$, where $f = a\cos(\varphi)$. It follows that F is in distributional sense equal to $F(x) = a(\langle x, d\rangle)\cos(\varphi\langle x, d\rangle)$, where $A := \mathscr{L}_d(a) = a(\langle\ \cdot\ , d\rangle)$ is the amplitude and $\Phi = \varphi(\langle\ \cdot\ , d\rangle)$ is the phase.

Chapter 3

Clifford algebras and Clifford modules

This chapter gives an introduction into Clifford algebras which is needed to understand the concepts used in the following chapters. Section 3.1 states basic facts about Clifford algebras which are needed throughout the rest of this work. Section 3.2 introduces Clifford modules to give the setting in which we then introduce some already known facts about functional analysis in Clifford Hilbert modules. Finally in section 3.2.3 we will for the first time systematically consider left-linear operators on Clifford-Hilbert modules. This results in new insights stated in Theorem 3.2.7 – Corollary 3.2.13. These insights will be essential for deriving uncertainty relations in chapter 6 and the development of Clifford frames in section 5.2.

The final part of the chapter treats some basics of Clifford analysis which will be useful in chapter 4.

3.1 Clifford algebra

In Theorem 2.1.7 we saw that the name analytical signal is due to a connection to complex analysis. Thus, a complement to complex numbers and complex analysis for general dimension would be desirable.

Generalisations of complex numbers are provided by Clifford algebras, which we will introduce now. The contents of this section is common knowledge about Clifford algebras found for example in [21] and [19].

3.1.1 Definition of Clifford algebras

The complex numbers $\mathbb{C}$ are $\mathbb{R}^2$ equipped with an algebraic structure. We will now define an algebraic structure (i.e. a multiplication) on $\mathbb{R}^n$.

Definition 3.1.1 (Clifford algebra)
LET $n \in \mathbb{N}$ *and* $\{e_0, e_1, \ldots, e_n\}$ *be an OrthoNormal Basis (ONB) of* $\mathbb{R}^{n+1}$.

THEN *a multiplication* $(a, b) \to ab$ *on* $\mathbb{R}^n \times \mathbb{R}^n$ *is defined by its action on the basis elements*

(i) $e_\alpha e_\alpha = -e_0$

(ii) $e_\alpha e_\beta = -e_\beta e_\alpha,\ \forall \alpha, \beta \in \{1,\ldots,n\} : \alpha \neq \beta$.

The resulting algebra over $\mathbb{R}^n$ *respectively* $\mathbb{R}^{n+1}$ *is called* **Clifford algebra** *and denoted by* $\mathbb{R}_n$.

Note that for every element $\{\alpha_1,\ldots,\alpha_\nu\}$ of the power set $\mathscr{P}(\{1,\ldots,n\})$, with $\nu \in \{1,\ldots,n\}$ elements, there exists by this definition an element $e_{\{\alpha_1,\ldots,\alpha_\nu\}} := e_{\alpha_1}\cdots e_{\alpha_\nu}$. Linear combinations of such elements are called ν-**vectors**. For the empty set $\emptyset$ the corresponding element is denoted by $e_0 := e_\emptyset$. It follows that e_0 is the unit element, i.e.,

$$\begin{aligned} &e_0 e_\alpha = e_\alpha e_0\ \forall \alpha \in \{1,\ldots,n\}, \text{ and} \\ &e_0 e_0 = e_0. \end{aligned}$$

This definition is clearly inspired by the construction of complex numbers. Indeed examples for Clifford algebras are $\mathbb{R}_0 = \mathbb{R}$, $\mathbb{R}_1 = \mathbb{C}$ and $\mathbb{R}_2 = \mathbb{H}$, (the Quaternions) which are the associative division algebras. However $\mathbb{R}_3 \neq \mathbb{O}$ (the Octonions) which can be easily seen, as Clifford algebras are associative, whereas Octonions are not.

It should be noted that in a more general definition of Clifford algebra, item (i) in our definition is replaced by a more general quadratic form. We will see later that our definition allows us to define a Clifford analysis without encountering zero divisors.

3.1.2 Properties of Clifford algebras

We know that vectors are elements of a Clifford algebra. Let us now have a look at the general elements of this algebra. To simplify notation we will use multivectors $\boldsymbol{\alpha}_\nu = \{\alpha_1,\ldots,\alpha_\nu\}$, where $\nu \in \{1,\ldots,n\} \cup \{0\}$ is the number of the elements $\alpha_k \in \{1,\ldots,n\}$. We will often skip the number of elements, simply writing $\boldsymbol{\alpha}_\nu = \boldsymbol{\alpha}$. The set of ordered elements of the product set $\mathscr{P}(\{1,\ldots,n\})$ play a special role hence we will denote them as

$$\mathscr{O}_n := \big\{\boldsymbol{\alpha}_\nu \in \mathscr{P}(\{1,\ldots,n\}) : 1 \leq \alpha_1 < \alpha_2 < \ldots < \alpha_\nu \leq n;\ \nu \in \{1,\ldots,n\} \cup \{0\}\big\}. \tag{3.1}$$

Let $\boldsymbol{\alpha}, \boldsymbol{\beta} \in \mathscr{O}_n$ then there exists $s \in \{-1, 1\}$ and $\boldsymbol{\gamma} \in \mathscr{O}_n$ such that $e_{\boldsymbol{\gamma}} = s e_{\boldsymbol{\alpha}} e_{\boldsymbol{\beta}}$. We will henceforth denote $\operatorname{sgn}(\boldsymbol{\alpha}\boldsymbol{\beta}) := s$.

The elements $e_{\boldsymbol{\alpha}_\nu}$ may be visualised as oriented areas when $\nu = 2$ and oriented volumes when $\nu = 3$. By Definition 3.1.1(ii) the elements corresponding to the elements of the ordered subset $\mathscr{O}_n$ of the power set give a basis for the algebra regarded as a linear space. This basis consists of 2^n elements.

Hence a general element of a Clifford algebra can be written in this basis as

$$x = \sum_{\substack{\boldsymbol{\alpha}_\nu \in \mathscr{O}_n \\ \nu \in \{0,\ldots,n\}}} x_{\boldsymbol{\alpha}_\nu} e_{\boldsymbol{\alpha}_\nu} = \sum_{\boldsymbol{\alpha} \in \mathscr{O}_n} x_{\boldsymbol{\alpha}} e_{\boldsymbol{\alpha}};\ x_{\boldsymbol{\alpha}} \in \mathbb{R}.$$

We denote the part of an element of the Clifford algebra corresponding to the basis element $e_{\boldsymbol{\beta}}$ ($\boldsymbol{\beta} \in \mathscr{O}_n$) by

$$\langle x \rangle_{\boldsymbol{\beta}} = \Big\langle \sum_{\boldsymbol{\alpha} \in \mathscr{O}_n} x_{\boldsymbol{\alpha}} e_{\boldsymbol{\alpha}} \Big\rangle_{\boldsymbol{\beta}} = x_{\boldsymbol{\beta}} \in \mathbb{R}.$$

The element corresponding to the empty set e_0 plays a special role, since $e_0 e_\alpha = e_\alpha = e_\alpha e_0$ and hence $e_0 e_0 = e_0$, it holds true that $a e_0 x = \sum_{\alpha \in \mathscr{O}_n} a x_\alpha e_\alpha$. Thus, e_0 is the identity of the algebra, and multiplication with elements of $\operatorname{span}\{e_0\}$ is equal to scalar multiplication. Hence, $\operatorname{span}\{e_0\}$ may be identified with $\mathbb{R}$.

3.1.3 Zero divisors

The Clifford algebras $\mathbb{R}_0 = \mathbb{R}$, $\mathbb{R}_1 = \mathbb{C}$ and $\mathbb{R}_2 = \mathbb{H}$ are division algebras and thus posses no zero divisors.

EXAMPLE 3.1.1 (Zero divisors):
For $n \geq 3$, the element $\frac{1}{2}(1 + e_{123})$ is an idempotent, and the corresponding zero divisor is $(1 + e_{123})(1 - e_{123}) = 0$.

Proof. To see this let us first compute e_{123}^2:

$$e_{123}e_{123} = -e_{123}e_1e_3e_2 = e_{12}e_3e_3e_{12} = e_{12}e_{21} = -e_1^2 = 1.$$

Thus we see that

$$\frac{1}{4}(1 + e_{123})(1 + e_{123}) = \frac{1}{4}(1 + 2e_{123} + e_{123}^2) = \frac{1}{2}(1 + e_{123}),$$

respectively that

$$(1 + e_{123})(1 - e_{123}) = 1 - e_{123} + e_{123} - e_{123}^2 = 0.$$

□

3.1.4 Paravectors and the paravector group

Definition 3.1.2 (Paravectors and paravector group)
LET $n \in \mathbb{N}$, *and* $\mathbb{R}_n$ *the corresponding Clifford algebra.*

THEN *elements of the form*

$$a_0 + \vec{a} := \sum_{\alpha=0}^{n} a_\alpha e_\alpha$$

are called **paravectors**. a_0 *is called the* **real part** *or* **scalar part**, *and* $\vec{a}$ *the* **vector part** *of the paravector. The paravectors form the* $n+1$*-dimensional linear subspace* $\mathbb{R}^{n+1}$ *of the Cliffordalgebra* $\mathbb{R}_n$.

The group $\{a \in \mathbb{R}_n : a = \prod_l a_l,\ a_l \in \mathbb{R}^{n+1}\}$ *spanned by the paravectors under multiplication is called the* **paravector group**.

Theorem 3.1.3 (The paravector group and zero deivisors)
LET $n \in \mathbb{N}$.

THEN *the paravector group of* $\mathbb{R}_n$ *is free of zero divisors.*

Proof. For a proof, see [19] Corollary (5.20). □

3.1.5 Scalar product, Euclidean measure and conjugation

Definition 3.1.4 (A scalar product on $\mathbb{R}_n$)
The Euclidean scalar product of the embedding of $\mathbb{R}_n$ *into* $\mathbb{R}^{2^n}$ *defines a* **scalar product** $(\cdot,\cdot)$ *on* $\mathbb{R}_n$*:*

$$(a,b) = \left(\sum_{\boldsymbol{\alpha}\in\mathcal{O}_n} a_{\boldsymbol{\alpha}} e_{\boldsymbol{\alpha}}, \sum_{\boldsymbol{\alpha}\in\mathcal{O}_n} b_{\boldsymbol{\alpha}} e_{\boldsymbol{\alpha}}\right) := \sum_{\boldsymbol{\alpha}\in\mathcal{O}_n} a_{\boldsymbol{\alpha}} b_{\boldsymbol{\alpha}}$$

This scalar product induces a norm $|\cdot| = (\cdot,\cdot)^{1/2}$ *which is called* **modulus**.

Corollary 3.1.5
$\big(\mathbb{R}_n, (\cdot,\cdot)\big)$ *is a real Hilbert space.*

Theorem 3.1.6 (Conjugation of paravectors)
A **conjugation** *is defined on the paravectors by*

$$\bar{\ }: \mathbb{R}^{n+1} \to \mathbb{R}^{n+1};\ x = x_0 + \sum_{\alpha=1}^{n} e_\alpha x_\alpha \mapsto x_0 - \sum_{\alpha=1}^{n} e_\alpha x_\alpha =: \overline{x},$$

This conjugation fulfills

$$(x,y) = \sum_{\alpha=0}^{n} x_\alpha y_\alpha = \langle x\overline{y}\rangle_0 = \frac{1}{2}(x\overline{y} + \overline{y}x),\ \forall x,y \in \mathbb{R}^{n+1}$$

and

$$(x,x) = x\overline{x} = |x|^2,\ \forall x \in \mathbb{R}^{n+1}.$$

As a consequence $(x,\overline{y}) = (\overline{x},y),\ \forall x,y \in \mathbb{R}^{n+1}$.

Proof. Let $x,y \in \mathbb{R}^{n+1}$.

$$\begin{aligned} x\overline{y} &= x_0y_0 - x_0\sum_\alpha e_\alpha y_\alpha + y_0\sum_\alpha e_\alpha x_\alpha - \sum_{\alpha,\beta} e_{\alpha\beta}x_\alpha y_\beta \\ &= x_0y_0 - x_0\sum_\alpha e_\alpha y_\alpha + y_0\sum_\alpha e_\alpha x_\alpha + \sum_\alpha x_\alpha y_\alpha - \sum_{\alpha<\beta} e_{\alpha\beta}(x_\beta y_\alpha - x_\alpha y_\beta) \end{aligned}$$

and

$$\begin{aligned} \overline{x}y &= x_0y_0 + x_0\sum_\alpha e_\alpha y_\alpha - y_0\sum_\alpha e_\alpha x_\alpha - \sum_{\alpha,\beta} e_{\alpha\beta}x_\alpha y_\beta \\ &= x_0y_0 + x_0\sum_\alpha e_\alpha y_\alpha - y_0\sum_\alpha e_\alpha x_\alpha + \sum_\alpha x_\alpha y_\alpha - \sum_{\alpha<\beta} e_{\alpha\beta}(x_\beta y_\alpha - x_\alpha y_\beta). \end{aligned}$$

Hence

$$\begin{aligned} x\overline{y} + y\overline{x} &= 2x_0y_0 - \sum_{\alpha<\beta} e_{\alpha\beta}\big((x_\beta y_\alpha - x_\alpha y_\beta) - (y_\beta x_\alpha - y_\alpha x_\beta)\big) + 2\sum_\alpha x_\alpha y_\alpha \\ &= 2(x,y). \end{aligned}$$

□

Corollary 3.1.7 (Inverses of paravectors)
LET $x \in \mathbb{R}^{n+1}$ *be a paravector.*
THEN x *is invertible with inverse* $\frac{\overline{x}}{|x|^2}$.

Proof. Note that $x\frac{\overline{x}}{|x|^2} = \frac{(x,x)}{|x|^2} = 1$. □

Definition 3.1.8 (Hypercomplex Conjugation)
The **hypercomplex conjugation** *on arbitrary elements of the Clifford algebra is defined by*

$$\begin{aligned} \overline{e}_\alpha &= -e_\alpha && \forall \alpha \in \{1,\dots,n\} \\ \overline{ab} &= \overline{b}\,\overline{a} && \forall a,b \in \mathbb{R}_n \end{aligned}$$

For the ν-vectors $e_{\boldsymbol{\alpha}_\nu}$, $\boldsymbol{\alpha}_\nu \in \mathcal{O}_n$, $\nu \in \{1,\dots,n\}$ of the canonical basis the conjugate is the inverse:

$$e_{\alpha_1,\dots,\alpha_\nu}\overline{e_{\alpha_1,\dots,\alpha_\nu}} = e_{\alpha_1,\dots,\alpha_{\nu-1}}e_{\alpha_\nu}(-e_{\alpha_\nu})\overline{e_{\alpha_1,\dots,\alpha_{\nu-1}}} = \dots = 1 \quad \forall \nu \in \{1,\dots,n\}.$$

Corollary 3.1.9
LET $e_{\boldsymbol{\alpha}_\nu}$, $\boldsymbol{\alpha}_\nu \in \mathscr{P}(\{1,\dots,n\})$, $\nu \le n$.
THEN *the conjugation of* $e_{\boldsymbol{\alpha}_\nu}$ *is equal to a change in sign:*

$$\overline{e_{\boldsymbol{\alpha}_\nu}} = (-1)^{\frac{\nu(\nu+1)}{2}} e_{\boldsymbol{\alpha}_\nu}.$$

As a consequence $\overline{\overline{e_{\boldsymbol{\alpha}_\nu}}} = e_{\boldsymbol{\alpha}_\nu}$.

Proof. $\overline{e_{\boldsymbol{\alpha}_\nu}} = \overline{e_{\alpha_\nu}}\dots\overline{e_{\alpha_1}} = (-1)^\nu e_{\alpha_\nu}\dots e_{\alpha_1} = (-1)^{\frac{\nu(\nu-1)}{2}}(-1)^\nu e_{\boldsymbol{\alpha}_\nu}$. □

Using the conjugation we define a $\mathbb{R}_n$-valued product.

Definition 3.1.10 (The $\mathbb{R}_n$-valued product $\langle\!\langle\ \cdot\ ,\ \cdot\ \rangle\!\rangle$)
The $\mathbb{R}_n$*-valued product* $\langle\!\langle\ \cdot\ ,\ \cdot\ \rangle\!\rangle$ *is defined by*

$$\langle\!\langle \cdot,\cdot \rangle\!\rangle : \mathbb{R}_n \times \mathbb{R}_n \to \mathbb{R}_n,\ \langle\!\langle x, y \rangle\!\rangle := x\overline{y}.$$

Corollary 3.1.11
LET $a \in \mathbb{R}^{n+1}$ *be an element of the paravector group.*
THEN $a\overline{a} = \langle\!\langle a, a \rangle\!\rangle = (a,a) = |a|^2$. *Furthermore a is invertible with the inverse* $\frac{\overline{a}}{(a,a)}$.

Proof. We know that this is true for paravectors. Thus for $a = \prod_{k=1}^{l} a_k \neq 0$, where $a_k \neq 0$ are paravectors, and

$$\langle\!\langle a, a \rangle\!\rangle = a\overline{a} = a_1 \dots \underbrace{a_l\overline{a_l}}_{\in\mathbb{R}^+} \dots \overline{a_1} = \prod_{k=1}^{l}(a_k, a_k) \in \mathbb{R}^+.$$

It follows that $\langle\!\langle a, a \rangle\!\rangle = (a,a)$. □

3.1.6 Complex Clifford algebras

We will consider the **complexification** $\mathbb{C}_n \cong \mathbb{C} \otimes \mathbb{R}_n$ of the Clifford algebra $\mathbb{R}_n$ given by

$$z = a + ib = \sum_{\boldsymbol{\alpha}\in\mathcal{O}_n} a_{\boldsymbol{\alpha}} e_{\boldsymbol{\alpha}} + i \sum_{\boldsymbol{\alpha}\in\mathcal{O}_n} b_{\boldsymbol{\alpha}} e_{\boldsymbol{\alpha}} = \sum_{\boldsymbol{\alpha}\in\mathcal{O}_n} z_{\boldsymbol{\alpha}} e_{\boldsymbol{\alpha}},\ z_{\boldsymbol{\alpha}} \in \mathbb{C},\ a_{\boldsymbol{\alpha}}, b_{\boldsymbol{\alpha}} \in \mathbb{R}.$$

The real part of an element of a **complex Clifford algebra** will be defined as $\Re z = \Re(\langle z \rangle_0)$.
Note that $\frac{1}{2}(1 \pm ie_\alpha)$ is an idempotent for all $\alpha \in \{1,\dots,n\}$, and hence a zero divisor. (Note that $(ie_\alpha)^2 = 1$).
Conjugation is defined as the combination of complex and hyper-complex conjugations

$$\overline{e_\alpha} = -e_\alpha,\ \forall \alpha \in \{1,\dots,n\}, \quad \overline{i} = -i \text{ and } \overline{ab} = \overline{b}\,\overline{a},\ \forall a, b \in \mathbb{C}_n.$$

A scalar product $(\cdot,\cdot)$ is defined by the Euclidean scalar product via the embedding of $\mathbb{C}_n$ into $\mathbb{R}^{2^{n+1}}$.

Corollary 3.1.12
LET $z = a + ib$, $x = l + ir$, *where* $a, b, l, r \in \mathbb{R}_n$.
THEN

(i) $(x, z) = \Re(x\overline{z})$.

(ii) $(xy, z) = (x, z\overline{y}) \quad \forall y \in \mathbb{C}_n$.

Proof. ad(i)

$$\Re(x\overline{z}) = \Big\langle \sum_{\alpha,\beta\in\mathcal{O}_n} e_\alpha\overline{e_\beta}\Re(x_\alpha\overline{z_\beta})\Big\rangle_0 = \Big\langle \sum_{\alpha,\beta\in\mathcal{O}_n} e_\alpha\overline{e_\beta}(l_\alpha a_\beta + r_\alpha b_\beta)\Big\rangle_0 = \sum_{\alpha\in\mathcal{O}_n} a_\alpha l_\alpha + b_\alpha r_\alpha$$
$$= (x, z).$$

ad(ii) $(xy, z) = \Re(xy\overline{z}) = \Re(x\overline{\overline{y}}\,\overline{z}) = \Re(x\overline{z\overline{y}}) = (x, z\overline{y})$.

□

3.2 Hypercomplex functional analysis

3.2.1 Clifford modules

We will now define Clifford modules especially over Hilbert spaces. The ultimate goal here is to give a setting in which frames on Clifford modules (especially on the module over $L^2(\mathbb{R}^n, \mathbb{R}_n)$) may be defined.

The first two sections are based on the according results in [32] and [7]. Theorem 3.2.4 is a corrected and expanded version of [32] Proposition 1.9.

The final section 3.2.3 consists of novel facts on left-linear operators on Clifford-Hilbert modules.

Definition 3.2.1 (Clifford module)
LET V *be a* $\mathbb{K}$*-vector space, where* $\mathbb{K} = \mathbb{R}$ *or* $\mathbb{K} = \mathbb{C}$.
THEN *the corresponding* **two-sided $\mathbb{K}$ Clifford module** *is defined as*

$$V_n := V \otimes \mathbb{R}_n := \left\{ x = \sum_{\alpha\in\mathcal{O}_n} x_\alpha \otimes e_\alpha; \quad x_\alpha \in V \right\}$$

and $\mathbb{K}_n$ *acts on* V_n *by*

$$ax := \sum_{\alpha,\beta\in\mathcal{O}_n} a_\beta x_\alpha \otimes e_\beta e_\alpha, \quad xa := \sum_{\alpha,\beta\in\mathcal{O}_n} a_\beta x_\alpha \otimes e_\alpha e_\beta; \quad \forall x \in V_n, a \in \mathbb{K}_n.$$

We will define the $\boldsymbol{\alpha}$*-part of an element* $x \in V_n$ *to be*

$$\langle x\rangle_\alpha = \Big\langle \sum_{\beta\in\mathcal{O}_n} x_\beta \otimes e_\beta \Big\rangle_\alpha := x_\alpha.$$

The **real part** *of an* $x \in V_n$ *will be defined as*

$$\Re x = \Re x_0 \otimes e_0.$$

If $(V, \|\cdot\|_V)$ is a normed vector space then V_n can be equipped with the Euclidean norm

$$\left\| \sum_{\boldsymbol{\alpha}\in\mathcal{O}_n} x_{\boldsymbol{\alpha}} \otimes e_{\boldsymbol{\alpha}} \right\|_{V_n} := \left(\sum_{\boldsymbol{\alpha}\in\mathcal{O}_n} \|x_{\boldsymbol{\alpha}}\|_V^2 \right)^{1/2}.$$

If $W \subseteq V_n$ is a left-(right-) submodule of V_n then a left (right) $\mathbb{C}_n$-linear mapping

$$L : W \mapsto \mathbb{C}_n$$

is called a **Clifford functional** *of W. The collection of all bounded Clifford functionals of W is called the* **dual Clifford module** *and is denoted by W^*.*
Note that $l : V_n \mapsto \mathbb{C}_n$ is called **left-linear** *iff $l(ax) = al(x)$ (respectively right linear $l(xa) = l(x)a$ for right-modules).*

In the following we will often write $x = \sum_{\boldsymbol{\alpha}\in\mathcal{O}_n} x_{\boldsymbol{\alpha}} e_{\boldsymbol{\alpha}} := \sum_{\boldsymbol{\alpha}\in\mathcal{O}_n} x_{\boldsymbol{\alpha}} \otimes e_{\boldsymbol{\alpha}}$ for an element $x \in V_n$.

Proposition 3.2.2 (Clifford functionals and functionals)
LET V *be a complex normed vector space.*

(i) LET $l : V_n \mapsto \mathbb{C}$ *be a linear functional.*
THEN $\widetilde{l}(x) := \sum_{\boldsymbol{\alpha}\in\mathcal{O}_n} l(\overline{e_{\boldsymbol{\alpha}}}x)e_{\boldsymbol{\alpha}}$ *is a (left-linear) Clifford functional. Furthermore $l = \langle \widetilde{l} \rangle_0$. More generally, for every $\boldsymbol{\alpha} \in \mathcal{O}_n$ there exists a uniquely defined (left-)linear functional $l_{[\boldsymbol{\alpha}]}(x) = \widetilde{l}(e_{\boldsymbol{\alpha}}x)$ such that $l(x) = \langle l_{[\boldsymbol{\alpha}]}(x)\rangle_{\boldsymbol{\alpha}}$.*

(ii) LET $h : V_n \mapsto \mathbb{C}_n$ *be a Clifford functional.*
THEN $l = \langle h \rangle_0$ *is a $\mathbb{C}$-linear functional and $\widetilde{l} = h$.*

(iii) LET h *be a (left-)Clifford functional.*
THEN $h = 0 \Leftrightarrow \exists \boldsymbol{\alpha} \in \mathcal{O}_n : \langle h \rangle_{\boldsymbol{\alpha}} = 0$.

(iv) h is a bounded Clifford functional iff $\langle h \rangle_{\boldsymbol{\alpha}}$ is a bounded functional for some $\boldsymbol{\alpha} \in \mathcal{O}_n$.

Proof. For the proof we will follow [7].

ad(i) By construction it is clear that $\widetilde{l}$ is $\mathbb{C}$-linear and that $l = \langle \widetilde{l} \rangle_0$. We need to show that

$$\widetilde{l}(ax) = a\widetilde{l}(x),\ \forall a \in \mathbb{C}_n.$$

Let $a = \sum_{\boldsymbol{\beta}\in\mathcal{O}_n} a_{\boldsymbol{\beta}} e_{\boldsymbol{\beta}} = \sum_{\nu=0}^{n} \sum_{\substack{\boldsymbol{\beta}_\nu\in\mathcal{O}_n \\ |\boldsymbol{\beta}_\nu|=\nu}} a_{\boldsymbol{\beta}_\nu} e_{\boldsymbol{\beta}_\nu}$. Then

$$\begin{aligned}
\widetilde{l}(ax) &= \sum_{\boldsymbol{\alpha}\in\mathcal{O}_n} l(\overline{e_{\boldsymbol{\alpha}}}ax)e_{\boldsymbol{\alpha}} = \sum_{\boldsymbol{\alpha},\boldsymbol{\beta}\in\mathcal{O}_n} a_{\boldsymbol{\beta}} l(\overline{e_{\boldsymbol{\alpha}}}e_{\boldsymbol{\beta}}x)e_{\boldsymbol{\alpha}} \\
&\overset{*}{=} \sum_{\nu=0}^{n} \sum_{\boldsymbol{\alpha},\boldsymbol{\beta}_\nu\in\mathcal{O}_n} (-1)^{\frac{\nu(\nu+1)}{2}} a_{\boldsymbol{\beta}_\nu} l(\overline{e_{\boldsymbol{\beta}_\nu}e_{\boldsymbol{\alpha}}}x)\overline{e_{\boldsymbol{\beta}_\nu}}e_{\boldsymbol{\beta}_\nu}e_{\boldsymbol{\alpha}} \\
&\overset{\boldsymbol{\gamma}=\boldsymbol{\beta}_\nu\boldsymbol{\alpha}}{=} \sum_{\nu=0}^{n} \sum_{\boldsymbol{\gamma},\boldsymbol{\beta}_\nu\in\mathcal{O}_n} a_{\boldsymbol{\beta}_\nu} l(\overline{e_{\boldsymbol{\gamma}}}x)(-1)^{\frac{\nu(\nu+1)}{2}}\overline{e_{\boldsymbol{\beta}_\nu}}e_{\boldsymbol{\gamma}} \\
&= \sum_{\boldsymbol{\gamma},\boldsymbol{\beta}\in\mathcal{O}_n} a_{\boldsymbol{\beta}} l(\overline{e_{\boldsymbol{\gamma}}}x)e_{\boldsymbol{\beta}}e_{\boldsymbol{\gamma}} \\
&= a\widetilde{l}(x),\ \forall x \in V_n.
\end{aligned}$$

To see * note that $\boldsymbol{\beta}_\nu \in \left\{\boldsymbol{\beta} \in \mathscr{P}(\{1,\dots,n\}) : 1 \le \beta_1 < \beta_2 < \ldots < \beta_\nu \le n\right\} = \{\boldsymbol{\beta} \in \mathscr{O}_n : |\boldsymbol{\beta}| = \nu\}$. By Definition 3.1.8 and Corollary 3.1.9

$$\overline{e_{\boldsymbol{\alpha}}} e_{\boldsymbol{\beta}_\nu} = \overline{e_{\boldsymbol{\alpha}}} (-1)^{\frac{\nu(\nu+1)}{2}} \overline{e_{\boldsymbol{\beta}_\nu}} = (-1)^{\frac{\nu(\nu+1)}{2}} \overline{e_{\boldsymbol{\beta}_\nu} e_{\boldsymbol{\alpha}}}.$$

Furthermore if we set $l_{[\boldsymbol{\alpha}]}(x) := \widetilde{l}(e_{\boldsymbol{\alpha}} x) = \sum_{\boldsymbol{\beta} \in \mathscr{O}_n} e_{\boldsymbol{\beta}} l(e_{\boldsymbol{\alpha}} \overline{e_{\boldsymbol{\beta}}} x)$, then $l_{[\boldsymbol{\alpha}]}$ is $\mathbb{C}_n$-linear by construction. Hence $\langle \widetilde{l}(e_{\boldsymbol{\alpha}} x)\rangle_{\boldsymbol{\alpha}} = \langle e_{\boldsymbol{\alpha}} \widetilde{l}(x)\rangle_{\boldsymbol{\alpha}} = \langle \widetilde{l}(x)\rangle_0 = l(x)$.

We will now proof the uniqueness of $l_{[\boldsymbol{\alpha}]}$. For any $\boldsymbol{\alpha} \in \mathscr{O}_n$ and $a \in \mathbb{C}_n$ let the $\mathbb{C}$-linear functional $t_{a,\boldsymbol{\alpha}}$ be defined by

$$t_{a,\boldsymbol{\alpha}} : \mathbb{C}_n \mapsto \mathbb{C};\ b \mapsto t_{a,\boldsymbol{\alpha}}(b) := l(e_{\boldsymbol{\alpha}} b a).$$

Then by the Riesz representation theorem (see for example [56] V.3.6) on $\mathbb{C}_n$, interpreted as the complex Hilbert space $\left(\mathbb{C}^{2^n}, (\,\cdot\,,\,\cdot\,)\right)$, there exists a unique $\mu_{a,\boldsymbol{\alpha}} \in \mathbb{C}_n$ such that

$$t_{a,\boldsymbol{\alpha}}(b) = (b, \mu_{a,\boldsymbol{\alpha}}).$$

Putting $b = \overline{e_{\boldsymbol{\alpha}}}$ we obtain

$$\langle \overline{\mu_{a,\boldsymbol{\alpha}}} \rangle_{\boldsymbol{\alpha}} = (\overline{e_{\boldsymbol{\alpha}}}, \mu_{a,\boldsymbol{\alpha}}) = t_{a,\boldsymbol{\alpha}}(\overline{e_{\boldsymbol{\alpha}}}) = l(e_{\boldsymbol{\alpha}} \overline{e_{\boldsymbol{\alpha}}} a) = l(a).$$

If we define

$$\lambda_{[\boldsymbol{\alpha}]}(a) := \overline{\mu_{a,\boldsymbol{\alpha}}},$$

then

$$\langle \lambda_{[\boldsymbol{\alpha}]}(a) \rangle_{\boldsymbol{\alpha}} = l(a).$$

Furthermore $\lambda_{[\boldsymbol{\alpha}]}$ is $\mathbb{C}_n$-linear as seen by the definition using Corollary 3.1.12: Let $c, d \in \mathbb{C}_n$.

$$\begin{aligned}(b, \overline{\lambda_{[\boldsymbol{\alpha}]}(c + da)}) &= (b, \mu_{c+da,\boldsymbol{\alpha}}) = t_{c+da,\boldsymbol{\alpha}}(b) = l(e_{\boldsymbol{\alpha}} b(c + da)) \\ &= l(e_{\boldsymbol{\alpha}} bc) + l(e_{\boldsymbol{\alpha}}(bd)a) = (b, \overline{\lambda_{[\boldsymbol{\alpha}]}(c)}) + (bd, \overline{\lambda_{[\boldsymbol{\alpha}]}(a)}) \\ &= (b, \overline{\lambda_{[\boldsymbol{\alpha}]}(c) + d\lambda_{[\boldsymbol{\alpha}]}(a)}).\end{aligned}$$

Now $\widetilde{l}(e_{\boldsymbol{\alpha}} a) = \lambda_{[\boldsymbol{\alpha}]}(a)$ since

$$\begin{aligned}\left(b, l_{[\boldsymbol{\alpha}]}(a)\right) &= \left\langle b \sum_{\boldsymbol{\beta} \in \mathscr{O}_n} \overline{e_{\boldsymbol{\beta}}} \langle l_{[\boldsymbol{\alpha}]}(a) \rangle_{\boldsymbol{\beta}} \right\rangle_0 = \sum_{\boldsymbol{\beta} \in \mathscr{O}_n} b_{\boldsymbol{\beta}} l(e_{\boldsymbol{\alpha}} \overline{e_{\boldsymbol{\beta}}} a) \\ &= \overline{l(e_{\boldsymbol{\alpha}} \overline{b} a)} = \overline{t_{a,\boldsymbol{\alpha}}(\overline{b})} = \overline{(\overline{b}, \mu_{a,\boldsymbol{\alpha}})} = (b, \overline{\mu_{a,\boldsymbol{\alpha}}}) \\ &= \left(b, \lambda_{[\boldsymbol{\alpha}]}(a)\right) \quad \forall b \in \mathbb{C}_n,\end{aligned}$$

whence $l_{[\boldsymbol{\alpha}]}$ is unique.

ad(ii) We will first show that l is $\mathbb{C}$-linear. Let $z \in \mathbb{C} \cong \mathbb{C}\{e_0\}$ then

$$l(zx) = \langle h(zx) \rangle_0 = \langle z h(x) \rangle_0 = z l(x),\ \forall x \in V_n.$$

Now

$$\begin{aligned}\widetilde{l}(x) &= \sum_{\boldsymbol{\alpha} \in \mathscr{O}_n} l(\overline{e_{\boldsymbol{\alpha}}} x) e_{\boldsymbol{\alpha}} = \sum_{\boldsymbol{\alpha} \in \mathscr{O}_n} \langle h(\overline{e_{\boldsymbol{\alpha}}} x) \rangle_0 e_{\boldsymbol{\alpha}} = \sum_{\boldsymbol{\alpha} \in \mathscr{O}_n} \langle \overline{e_{\boldsymbol{\alpha}}} h(x) \rangle_0 e_{\boldsymbol{\alpha}} \\ &= \sum_{\boldsymbol{\alpha} \in \mathscr{O}_n} \left\langle \sum_{\boldsymbol{\beta} \in \mathscr{O}_n} \overline{e_{\boldsymbol{\alpha}}} e_{\boldsymbol{\beta}} \langle h(x) \rangle_{\boldsymbol{\beta}} \right\rangle_0 e_{\boldsymbol{\alpha}} = \sum_{\boldsymbol{\alpha} \in \mathscr{O}_n} \langle h(x) \rangle_{\boldsymbol{\alpha}} e_{\boldsymbol{\alpha}} \\ &= h(x),\ \forall x \in V_n.\end{aligned}$$

ad(iii) Let h be a Clifford functional. If $h = 0$ then $\langle h\rangle_{\boldsymbol{\alpha}} = 0$, $\forall \boldsymbol{\alpha} \in \mathscr{O}_n$.

Conversely suppose $h_{\boldsymbol{\alpha}} = 0$ for some $\boldsymbol{\alpha} \in \mathscr{O}_n$. Then h is the unique Clifford functional such that $\big(\langle h\rangle_{\boldsymbol{\alpha}}\big)_{[\boldsymbol{\alpha}]} = h$. Hence by the proof of Proposition 3.2.2(i)

$$h(a) = \sum_{\boldsymbol{\beta}\in\mathscr{O}_n} e_{\boldsymbol{\beta}}\langle h(\overline{e_{\boldsymbol{\beta}}}a)\rangle_0 = \sum_{\boldsymbol{\beta}\in\mathscr{O}_n} e_{\boldsymbol{\beta}}\langle h(e_{\boldsymbol{\alpha}}\overline{e_{\boldsymbol{\beta}}}a)\rangle_{\boldsymbol{\alpha}} = 0.$$

ad(iv) Let $h \in V_n^*$. Then, $\langle h\rangle_{\boldsymbol{\alpha}}$ is bounded for all $\boldsymbol{\alpha} \in \mathscr{O}_n$. Conversely let $\boldsymbol{\alpha} \in \mathscr{O}_n$ and let

$$|\langle h(a)\rangle_{\boldsymbol{\alpha}}| \le C|a|,\ \forall a \in V_n.$$

Then

$$\begin{aligned}|h(a)|^2 &= \sum_{\boldsymbol{\beta}\in\mathscr{O}_n} \left|\left\langle h(a)\right\rangle_{\boldsymbol{\beta}}\right|^2 = \sum_{\boldsymbol{\beta}\in\mathscr{O}_n} \left|\langle h(\overline{e_{\boldsymbol{\beta}}}a)\rangle_0\right|^2 = \sum_{\boldsymbol{\beta}\in\mathscr{O}_n} \left|\left\langle h(e_{\boldsymbol{\alpha}}\overline{e_{\boldsymbol{\beta}}}a)\right\rangle_{\boldsymbol{\alpha}}\right|^2 \\ &\le \sum_{\boldsymbol{\beta}\in\mathscr{O}_n} C^2|e_{\boldsymbol{\alpha}}\overline{e_{\boldsymbol{\beta}}}a|^2 \le \sum_{\boldsymbol{\beta}\in\mathscr{O}_n} C^2|e_{\boldsymbol{\alpha}}|^2|\overline{e_{\boldsymbol{\beta}}}a|^2 \\ &\le 2^n C^2|a|^2.\end{aligned}$$

□

Theorem 3.2.3 (The Hahn-Banach Theorem)
LET V *be a complex normed vector space and let* X_l *be a left-submodule of* V_n*. Furthermore let* h *be a left-linear Clifford functional on* X_l*.*
THEN *there exists a left-linear Clifford functional* h^* *on* V_n *such that*

$$h^*\big|_{X_l} = h \text{ and } \|h\| \le \|h^*\| \le 2^{n/2}\|h\|.$$

Proof. The theorem was stated in [32] without a proof. Hence we will give the proof.

By the complex case of the Hahn-Banach theorem (see for example [56] III.1.5) we can extend $\langle h\rangle_0$ to a complex linear functional on V_n denoted by $\langle h^*\rangle_0$ of the same norm $\|\langle h\rangle_0\| = \|\langle h^*\rangle_0\|$.

By Proposition 3.2.2(ii) we know that $\widetilde{\langle h^*\rangle_0}\big|_{X_l} = h$. If we denote $h^* := \widetilde{\langle h^*\rangle_0}$, then $\|\langle h^*\rangle_0\| = \|\langle h\rangle_0\|$ is bounded, whence $\|h\| \le \|h^*\|$ and by Proposition 3.2.2(iv) $\|h^*\| \le \|\langle h\rangle_0\| \le 2^{n/2}\|h\|$.

□

3.2.2 Clifford-Hilbert modules

Let $\big(H, \langle\,\cdot\,,\,\cdot\,\rangle_H\big)$ be a complex Hilbert space. We will now consider $H_n := H \otimes \mathbb{C}_n$ the space, that has elements of the form $x = \sum_{\boldsymbol{\alpha}\in\mathscr{O}_n} x_{\boldsymbol{\alpha}} e_{\boldsymbol{\alpha}}$, where $x_{\boldsymbol{\alpha}} \in H$, $\forall \boldsymbol{\alpha} \in \mathscr{O}_n$. H_n is a complex Hilbert space when endowed with the scalar product

$$(\,\cdot\,,\,\cdot\,) : H_n \times H_n \to \mathbb{C},\ (x, y) := \sum_{\boldsymbol{\alpha}\in\mathscr{O}_n} \langle x_{\boldsymbol{\alpha}}, y_{\boldsymbol{\alpha}}\rangle_H,\ \forall x, y \in H_n.$$

This scalar product induces the Euclidean metric

$$\|x\|_{H_n}^2 := (x, x) = \sum_{\boldsymbol{\alpha}\in\mathscr{O}_n} \|x_{\boldsymbol{\alpha}}\|_H^2.$$

The Euclidean scalar product just introduced is not Clifford algebra-valued. Inspired by the scalar product on $\mathbb{C}$ we will now introduce a complexified Clifford algebra-valued form on H_n.

$$\langle\!\langle \cdot , \cdot \rangle\!\rangle : \mathbb{H}_n \times \mathbb{H}_n \to \mathbb{C}_n, \; \langle\!\langle x, y\rangle\!\rangle := \sum_{\boldsymbol{\alpha},\boldsymbol{\beta}\in\mathscr{O}_n} \langle x_{\boldsymbol{\alpha}}, y_{\boldsymbol{\beta}}\rangle_H e_{\boldsymbol{\alpha}} \overline{e_{\boldsymbol{\beta}}}$$

If $\overline{}$ is a conjugation on H, then

$$\overline{} : H_n \mapsto H_n, \quad x \mapsto \sum_{\boldsymbol{\alpha}\in\mathscr{O}_n} \overline{x_{\boldsymbol{\alpha}}}\,\overline{e_{\boldsymbol{\alpha}}}$$

defines a conjugation on H_n.

Theorem 3.2.4 (Properties of the Clifford algebra-valued inner product)
LET $x, y \in H_n$ *and* $a \in \mathbb{C}_n$.
THEN *the following hold.*

(i) $\langle\!\langle ax, y\rangle\!\rangle = a\langle\!\langle x, y\rangle\!\rangle$, $\quad \langle\!\langle x, ay\rangle\!\rangle = \langle\!\langle x, y\rangle\!\rangle\overline{a}$, $\quad \langle\!\langle xa, y\rangle\!\rangle = \langle\!\langle x, y\overline{a}\rangle\!\rangle$ *and* $\overline{\langle\!\langle x, y\rangle\!\rangle} = \langle\!\langle y, x\rangle\!\rangle$.

(ii) $\langle\!\langle x, y\rangle\!\rangle = \sum_{\boldsymbol{\alpha}\in\mathscr{O}_n} (\overline{e_{\boldsymbol{\alpha}}} x, y) e_{\boldsymbol{\alpha}}$, *in particular,* $\langle\, \langle\!\langle x, y\rangle\!\rangle\, \rangle_0 = (x, y)$.

(iii) $\left|\langle\!\langle x, y\rangle\!\rangle\right| \le 2^{n/2} \|x\| \|y\|$.

(iv) $\Re\langle\!\langle x, x\rangle\!\rangle = \|x\|^2$. *In particular,* $\|x\|^2 \le \left|\langle\!\langle x, x\rangle\!\rangle\right| \le 2^n \|x\|^2$.

(v) $\|ax\| \le 2^{n/2}|a|\|x\|$. *However if* $a \in \mathbb{R}^{n+1}$ *is a paravector, or an element of the paravector group, then* $\|ax\| = |a|\|x\|$.

(vi) $\overline{\overline{x}} = x$, $\quad \overline{ax} = \overline{x}\,\overline{a}$, $\quad \overline{xa} = \overline{a}\,\overline{x}$.

(vii) $\|x\| \le \sup_{\|y\|\le 1} \left|\langle\!\langle x, y\rangle\!\rangle\right| \le 2^{n/2}\|x\|$.

Proof. (i)–(vi) can be found in [32], while (vii) seems to be new.

ad(i)

$$\begin{aligned}
\langle\!\langle ax, y\rangle\!\rangle &= \sum_{\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\gamma}\in\mathscr{O}_n} \langle a_{\boldsymbol{\alpha}} x_{\boldsymbol{\beta}}, y_{\boldsymbol{\gamma}}\rangle_H e_{\boldsymbol{\alpha}} e_{\boldsymbol{\beta}} \overline{e_{\boldsymbol{\gamma}}} &&= \sum_{\boldsymbol{\alpha}\in\mathscr{O}_n} a_{\boldsymbol{\alpha}} e_{\boldsymbol{\alpha}} \sum_{\boldsymbol{\beta},\boldsymbol{\gamma}\in\mathscr{O}_n} \langle x_{\boldsymbol{\beta}}, y_{\boldsymbol{\gamma}}\rangle_H e_{\boldsymbol{\beta}} \overline{e_{\boldsymbol{\gamma}}} &&= a\langle\!\langle x, y\rangle\!\rangle.\\
\langle\!\langle x, ay\rangle\!\rangle &= \sum_{\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\gamma}\in\mathscr{O}_n} \langle x_{\boldsymbol{\alpha}}, a_{\boldsymbol{\beta}} y_{\boldsymbol{\gamma}}\rangle_H e_{\boldsymbol{\alpha}} \overline{e_{\boldsymbol{\beta}} e_{\boldsymbol{\gamma}}} &&= \sum_{\boldsymbol{\alpha},\boldsymbol{\gamma}\in\mathscr{O}_n} \langle x_{\boldsymbol{\alpha}}, y_{\boldsymbol{\gamma}}\rangle_H e_{\boldsymbol{\alpha}} \overline{e_{\boldsymbol{\gamma}}} \sum_{\boldsymbol{\beta}\in\mathscr{O}_n} \overline{a_{\boldsymbol{\beta}} e_{\boldsymbol{\beta}}} &&= \langle\!\langle x, y\rangle\!\rangle\overline{a}.\\
\langle\!\langle xa, y\rangle\!\rangle &= \sum_{\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\gamma}\in\mathscr{O}_n} \langle x_{\boldsymbol{\alpha}} a_{\boldsymbol{\beta}}, y_{\boldsymbol{\gamma}}\rangle_H e_{\boldsymbol{\alpha}} e_{\boldsymbol{\beta}} \overline{e_{\boldsymbol{\gamma}}} &&= \sum_{\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\gamma}\in\mathscr{O}_n} a_{\boldsymbol{\beta}} \langle x_{\boldsymbol{\alpha}}, y_{\boldsymbol{\gamma}}\rangle_H e_{\boldsymbol{\alpha}} \overline{\overline{e_{\boldsymbol{\beta}}}}\, \overline{e_{\boldsymbol{\gamma}}} &&\\
& &&= \sum_{\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\gamma}\in\mathscr{O}_n} \langle x_{\boldsymbol{\alpha}}, y_{\boldsymbol{\gamma}} \overline{a_{\boldsymbol{\beta}}}\rangle_H e_{\boldsymbol{\alpha}} \overline{e_{\boldsymbol{\gamma}} \overline{e_{\boldsymbol{\beta}}}} &&= \langle\!\langle x, y\overline{a}\rangle\!\rangle.\\
\overline{\langle\!\langle x, y\rangle\!\rangle} &= \sum_{\boldsymbol{\alpha},\boldsymbol{\beta}\in\mathscr{O}_n} \overline{\langle x_{\boldsymbol{\alpha}}, y_{\boldsymbol{\beta}}\rangle_H e_{\boldsymbol{\alpha}} \overline{e_{\boldsymbol{\beta}}}} &&= \sum_{\boldsymbol{\alpha},\boldsymbol{\beta}\in\mathscr{O}_n} \langle y_{\boldsymbol{\beta}}, x_{\boldsymbol{\alpha}}\rangle_H e_{\boldsymbol{\beta}} \overline{e_{\boldsymbol{\alpha}}} &&= \langle\!\langle y, x\rangle\!\rangle.
\end{aligned}$$

ad(ii) We easily see that by definition $(.,.) = \langle \langle\!\langle .,. \rangle\!\rangle \rangle_0$. Now (i) shows that $\langle\!\langle ., y\rangle\!\rangle$ suffices the requirements of Proposition 3.2.2(ii) which proves (ii).

ad(iii)

$$|\langle\!\langle x,y\rangle\!\rangle|^2 \overset{(ii)}{=} \Big|\sum_{\boldsymbol{\alpha}\in\mathcal{O}_n}\big(\overline{e_{\boldsymbol{\alpha}}}x,y\big)e_{\boldsymbol{\alpha}}\Big|^2 = \sum_{\boldsymbol{\alpha}\in\mathcal{O}_n}\big|\big(\overline{e_{\boldsymbol{\alpha}}x,y}\big)\big|^2 \overset{*}{\le} \sum_{\boldsymbol{\alpha}\in\mathcal{O}_n}\|\overline{e_{\boldsymbol{\alpha}}}x\|^2\|y\|^2$$
$$= 2^n\|x\|^2\|y\|^2,$$

where in $*$ the Cauchy-Schwartz-inequality in H has been used.

ad(iv) $\|x\|^2 = (x,x) = \langle\langle\!\langle x,x\rangle\!\rangle\rangle_0 = \Re\langle\!\langle x,x\rangle\!\rangle$. Hence $\|x\|^2 \le |\langle\!\langle x,x\rangle\!\rangle|$. Now $|\langle\!\langle x,x\rangle\!\rangle| \le 2^n\|x\|^2$ follows from (iii).

ad(v) Let $a \in \mathbb{R}^{n+1}$, then

$$\|xa\|^2 \overset{(iv)}{=} \Re\langle\!\langle xa,xa\rangle\!\rangle \overset{(i)}{=} \Re\langle\!\langle xa\overline{a},x\rangle\!\rangle \overset{*}{=} |a|^2\Re\langle\!\langle x,x\rangle\!\rangle = |a|^2\|x\|^2.$$

In $*$, we used the fact that for paravectors $a\overline{a} = |a|^2 \in \mathbb{R}$, hence the assumption $\langle a,a\rangle = |a|^2$ is sufficient for the equality to hold.

By definition it is clear that $|\overline{a}| = |a|$, $\forall a \in \mathbb{C}_n$, and $\|\overline{x}\| = \|x\|$, $\forall x \in H_n$. Hence $\|ax\| = \|\overline{\overline{x}\,\overline{a}}\| = \|\overline{x}\,\overline{a}\| = |\overline{a}|\|\overline{x}\| = |a|\|x\|$.

Let now $a \in \mathbb{C}_n$. Then

$$\|ax\|^2 = \Big\|\sum_{\boldsymbol{\alpha},\boldsymbol{\beta}\in\mathcal{O}_n} a_{\boldsymbol{\alpha}}x_{\boldsymbol{\beta}}e_{\boldsymbol{\alpha}}e_{\boldsymbol{\beta}}\Big\|^2 \le \Big(\sum_{\boldsymbol{\alpha}\in\mathcal{O}_n}|a_{\boldsymbol{\alpha}}|\Big\|\sum_{\boldsymbol{\beta}\in\mathcal{O}_n}x_{\boldsymbol{\beta}}e_{\boldsymbol{\alpha}}e_{\boldsymbol{\beta}}\Big\|\Big)^2$$
$$\overset{*}{\le} \sum_{\boldsymbol{\alpha}\in\mathcal{O}_n}|a_{\boldsymbol{\alpha}}|^2\sum_{\boldsymbol{\gamma}\in\mathcal{O}_n}\Big\|\sum_{\boldsymbol{\beta}\in\mathcal{O}_n}x_{\boldsymbol{\beta}}e_{\boldsymbol{\gamma}}e_{\boldsymbol{\beta}}\Big\|^2$$
$$= 2^n|a|^2\|x\|^2.$$

In * we used the Cauchy-Schwartz inequality.

ad(vi) $\overline{\overline{x}} = x$ is clear since this holds for the conjugation on H and $\mathbb{C}_n$. Now

$$\overline{ax} = \sum_{\boldsymbol{\alpha},\boldsymbol{\beta}\in\mathcal{O}_n}\overline{a_{\boldsymbol{\alpha}}x_{\boldsymbol{\beta}}e_{\boldsymbol{\alpha}}e_{\boldsymbol{\beta}}} = \sum_{\boldsymbol{\alpha},\boldsymbol{\beta}\in\mathcal{O}_n}\overline{x_{\boldsymbol{\beta}}}\,\overline{a_{\boldsymbol{\alpha}}}\,\overline{e_{\boldsymbol{\beta}}}\,\overline{e_{\boldsymbol{\alpha}}} = \overline{x}\,\overline{a}.$$

$\overline{xa} = \overline{a}\,\overline{x}$ follows in the same way.

ad(vii) Since $\big(H_n,(\cdot,\cdot)\big)$ is a complex Hilbert space

$$\|x\| = \sup_{\|y\|\le 1}|(x,y)| \le \sup_{\|y\|\le 1}|\langle\!\langle x,y\rangle\!\rangle|.$$

By (iii) $\sup_{\|y\|\le 1}|\langle\!\langle x,y\rangle\!\rangle| \le 2^{n/2}\|x\|$.

□

EXAMPLE 3.2.1 (Examples for Theorem 3.2.4(iii)):
We will now give examples for Theorem 3.2.4(iii) which demonstrate in which cases there is equality and in which cases the inequality holds without the factor $2^{n/2}$.

(i) LET $x,y \in H_n : (\overline{e_{\boldsymbol{\alpha}}}x,y) = \|x\|\|y\|$, $\forall \boldsymbol{\alpha} \in \mathcal{O}_n$.

THEN by Theorem 3.2.4(ii)

$$\big|\langle\!\langle x,y\rangle\!\rangle\big|^2 = \sum_{\boldsymbol{\alpha}\in\mathcal{O}_n}|e_{\boldsymbol{\alpha}}(\overline{e_{\boldsymbol{\alpha}}}x,y)|^2 = \sum_{\boldsymbol{\alpha}\in\mathcal{O}_n}\|x\|^2\|y\|^2 = 2^n\|x\|^2\|y\|^2.$$

(ii) LET $x, y \in H_n : \langle x_{\boldsymbol{\alpha}}, y_{\boldsymbol{\beta}}\rangle_H = 0,\ \forall \boldsymbol{\alpha} \neq \boldsymbol{\beta} \in \mathscr{O}_n$.

THEN

$$\left|\left\langle\!\left\langle x, y\right\rangle\!\right\rangle\right|^2 = \sum_{\boldsymbol{\alpha},\boldsymbol{\beta}\in\mathscr{O}_n} |\langle x_{\boldsymbol{\alpha}}, y_{\boldsymbol{\beta}}\rangle|^2 = \sum_{\boldsymbol{\alpha}\in\mathscr{O}_n} |\langle x_{\boldsymbol{\alpha}}, y_{\boldsymbol{\alpha}}\rangle|^2 \leq \sum_{\boldsymbol{\alpha}\in\mathscr{O}_n} \|x_{\boldsymbol{\alpha}}\|^2\|y_{\boldsymbol{\alpha}}\|^2 \leq \sum_{\boldsymbol{\alpha}\in\mathscr{O}_n} \|x_{\boldsymbol{\alpha}}\|^2 \sum_{\boldsymbol{\beta}} \|y_{\boldsymbol{\beta}}\|^2$$
$$= \|x\|^2\|y\|^2.$$

Theorem 3.2.5 (Riesz representation theorem)
LET V *be a left-submodule of* H_n *and* $L \in V^*$ *a Clifford functional.*

THEN *there exists a unique element* $a \in V$ *such that* $L(x) = \langle\!\langle x, a\rangle\!\rangle$*. Moreover* $\|a\| \approx \|L\|$*. The operator* $\Phi : V \mapsto V^*,\ a \mapsto \langle\!\langle \cdot, a\rangle\!\rangle$ *is bijective and conjugate linear.*

Proof. This theorem can be found in [32]. By the Riesz representation theorem for complex Hilbert spaces (see for example [56] V.3.6), $\exists a \in V : \langle L(.)\rangle_0 = (., a) = \langle\langle\!\langle ., a\rangle\!\rangle\rangle_0$. Let $x \in H_n$. Then

$$\begin{aligned} L(x) &= \sum_{\boldsymbol{\alpha}\in\mathscr{O}_n} \langle \overline{e_{\boldsymbol{\alpha}}} L(x)\rangle_0 e_{\boldsymbol{\alpha}} = \sum_{\boldsymbol{\alpha}\in\mathscr{O}_n} \langle L(\overline{e_{\boldsymbol{\alpha}}} x)\rangle_0 e_{\boldsymbol{\alpha}} \\ &= \sum_{\boldsymbol{\alpha}\in\mathscr{O}_n} \langle\langle\!\langle \overline{e_{\boldsymbol{\alpha}}} x, a\rangle\!\rangle\rangle_0 e_{\boldsymbol{\alpha}} = \sum_{\boldsymbol{\alpha}\in\mathscr{O}_n} \langle \overline{e_{\boldsymbol{\alpha}}}\langle\!\langle x, a\rangle\!\rangle\rangle_0 e_{\boldsymbol{\alpha}} \\ &= \langle\!\langle x, a\rangle\!\rangle. \end{aligned}$$

The representation is unique since $\langle\!\langle x, a\rangle\!\rangle = 0,\ \forall x \in V \Rightarrow \langle\langle\!\langle a, a\rangle\!\rangle\rangle_0 = \|a\| = 0$, and hence $a = 0$. Now the equivalence of the norms is seen by $\|a\|^2 = \langle\langle\!\langle a, a\rangle\!\rangle\rangle_0 \leq \left|\langle\!\langle a, a\rangle\!\rangle\right|$ whence

$$\|a\| \leq \left|\left\langle\!\left\langle \frac{a}{\|a\|}, a\right\rangle\!\right\rangle\right| \leq \sup_{\|y\|=1} |\langle\!\langle y, a\rangle\!\rangle| = \sup_{\|y\|=1} |L(y)| = \|L\|,$$

and $\sup_{\|y\|=1} |L(y)| = \sup_{\|y\|=1} |\langle\!\langle y, a\rangle\!\rangle| \overset{3.2.4(ii)}{\leq} 2^{n/2}\|a\|$. □

3.2.3 Operators on Clifford-Hilbert modules

In the following section we will consider left-linear operators on Clifford modules. It seems that this is the first time that these operators are considered systematically, whence all results have to be considered novel. However many results correspond quite closely to the results on linear operators on Hilbert spaces. The results which differ from the Hilbert space case are Theorem 3.2.7(ix), Corollary 3.2.8 and the results on self-adjoint operators Proposition 3.2.10 and Theorem 3.2.12.

Let $L(H_n, G_n)$ be the module of Clifford left-linear bounded operators from the Clifford-Hilbert module H_n to the Clifford-Hilbert module G_n.

For an operator $T \in L(H_n, G_n)$ an adjoint in the sense of Clifford modules of Banach spaces is defined by

$$T' : G_n^* \mapsto H_n^*,\ y^* \mapsto T'(y^*) : x \mapsto T'(y^*)(x) := y^*(Tx).$$

Definition 3.2.6 (The adjoint operator)
LET $T \in L(H_n, G_n)$.

THEN *its* **adjoint** *is defined by*

$$T^* : G_n \mapsto H_n;\ \langle\!\langle Tx, y\rangle\!\rangle_{G_n} = \langle\!\langle x, T^* y\rangle\!\rangle_{H_n}.$$

By Theorem 3.2.5 this is equal to $T^* = \Phi_{H_n}^{-1} T' \Phi_{G_n}$.

Theorem 3.2.7 (Operators and their adjoints)
LET H_n, G_n, F_n *be Clifford-Hilbert modules. Furthermore let* $S, T \in L(H_n, G_n)$*,*
$R \in L(G_n, F_n)$*, and* $a \in \mathbb{C}_n$*.*
THEN

(i) $(S+T)^* = S^* + T^*$.

(ii) $(Sa)^* = S^*\overline{a}$.

(iii) $(RS)^* = S^*R^*$.

(iv) $S^* \in L(G_n^*, H_n^*)$.

(v) $S^{**} = S$.

(vi) $\|S\| = \|S^*\|$.

(vii) $\|SS^*\| = \|S^*S\| = \|S\|^2$.

(viii) $\mathrm{Ker}(S) = \mathrm{Ran}(S^*)^\perp$; $\mathrm{Ker}(S^*) = \mathrm{Ran}(S)^\perp$.

(ix) *there exists a set of linear operators*

$$\{S_{\boldsymbol\beta}\}_{\boldsymbol\beta\in\mathscr{O}_n},\ S_{\boldsymbol\beta} : H \to G$$

such that

$$Sf = \sum_{\boldsymbol\alpha,\boldsymbol\beta\in\mathscr{O}_n} e_{\boldsymbol\alpha} e_{\boldsymbol\beta} S_{\boldsymbol\beta} f_{\boldsymbol\alpha}.$$

(Keep in mind that in order to be left-linear the operator $S = \sum_{\boldsymbol\alpha\in\mathscr{O}_n} e_{\boldsymbol\alpha} S_{\boldsymbol\alpha}$ *has to be applied from the right.)*

The operator S is bounded iff for every $\boldsymbol\alpha \in \mathscr{O}_n$ *the operator* $S_{\boldsymbol\alpha}$ *is bounded.*

Proof.

ad(i) $\langle\!\langle x, (S+T)^* y\rangle\!\rangle = \langle\!\langle (S+T)x, y\rangle\!\rangle = \langle\!\langle Sx, y\rangle\!\rangle + \langle\!\langle Tx, y\rangle\!\rangle = \langle\!\langle x, S^* y\rangle\!\rangle + \langle\!\langle x, T^* y\rangle\!\rangle$

ad(ii) $\langle\!\langle (Sa)x, y\rangle\!\rangle = \langle\!\langle Sxa, y\rangle\!\rangle = \langle\!\langle xa, S^* y\rangle\!\rangle \overset{3.2.4(v)}{=} \langle\!\langle x, S^* y\overline{a}\rangle\!\rangle = \langle\!\langle x, (S^*\overline{a}) y\rangle\!\rangle$

ad(iii) $\langle\!\langle RSx, y\rangle\!\rangle_{H_n} = \langle\!\langle Sx, R^* y\rangle\!\rangle_{G_n} = \langle\!\langle x, S^* R^* y\rangle\!\rangle_{F_n}$

ad(iv) We will show $S^* \in L(G_n^*, H_n^*)$.

S^* is Clifford linear, since $\langle\!\langle x, S^*(y+az)\rangle\!\rangle = \langle\!\langle Sx, y\rangle\!\rangle + \langle\!\langle Sx, z\rangle\!\rangle\overline{a} = \langle\!\langle x, S^* y + aS^* z\rangle\!\rangle$.

Boundedness follows since

$$\begin{aligned}\|S^*\|^2 &= \sup_{\|x\|=1} (S^*x, S^*x) = \langle\langle\!\langle x, SS^*x\rangle\!\rangle\rangle_0 = (x, SS^*x)\\ &\overset{*}{\le} \sup_{\|x\|=1} \|x\|\,\|SS^*x\| \le \sup_{\|x\|=1} \|S^*\|\,\|Sx\|\,\|x\|\\ &= \|S^*\|\,\|S\|,\end{aligned}$$

where $*$ follows from the Cauchy-Schwartz inequality on the Hilbert space $\big(H_n, (\cdot,\cdot)\big)$.

ad(v) By (iv) $S^{**} \in L(H_n^{**}, G_n^{**})$. Hence (v) is obvious, since

$$\langle\!\langle S^{**}x, y\rangle\!\rangle_{G_n^{**}} = \langle\!\langle x, S^* y\rangle\!\rangle_{H_n} = \langle\!\langle Sx, y\rangle\!\rangle_{G_n}$$

and $G_n^{**} = G_n$.

ad(vi) Using (iv) and (v) $\|S\| \geq \|S^*\| \geq \|S^{**}\| = \|S\|$.

ad(vii) Is obvious since it is true in the complex Hilbert space $\big(H_n, (\cdot,\cdot)\big)$.

ad(viii) Let $x \in \mathrm{Ker}(S)$. Then

$$\begin{aligned} Sx = 0 \quad & \Leftrightarrow \langle\!\langle Sx, y\rangle\!\rangle = 0 \ \forall y \in G_n \\ & \Leftrightarrow \langle\!\langle x, S^* y\rangle\!\rangle = 0 \ \forall y \in G_n \\ & \Leftrightarrow x \in \big(\mathrm{Ran}(S^*)\big)^{\perp}. \end{aligned}$$

Furthermore $\mathrm{Ker}(S^*) = \big(\mathrm{Ran}(S^{**})\big)^{\perp} = \big(\mathrm{Ran}(S)\big)^{\perp}$.

ad(ix) Let $S = \sum_{\boldsymbol{\alpha}\in\mathscr{O}_n} e_{\boldsymbol{\alpha}} S_{\boldsymbol{\alpha}} : H_n \to G_n,\ f = \sum_{\boldsymbol{\beta}\in\mathscr{O}_n} e_{\boldsymbol{\beta}} f_{\boldsymbol{\beta}} \mapsto \sum_{\boldsymbol{\alpha},\boldsymbol{\beta}\in\mathscr{O}_n} e_{\boldsymbol{\beta}} e_{\boldsymbol{\alpha}} S_{\boldsymbol{\alpha}} f_{\boldsymbol{\beta}}$. Furthermore let $a \in \mathbb{C}^n,\ f, g \in H_n$. Then

$$\begin{aligned} S(af+g) &= \sum_{\boldsymbol{\alpha},\boldsymbol{\beta}} e_{\boldsymbol{\beta}} e_{\boldsymbol{\alpha}} S_{\boldsymbol{\alpha}}(\langle af\rangle_{\boldsymbol{\beta}} + \langle g\rangle_{\boldsymbol{\beta}}) = \sum_{\boldsymbol{\alpha},\boldsymbol{\beta}} e_{\boldsymbol{\beta}} e_{\boldsymbol{\alpha}} \Big(S_{\boldsymbol{\alpha}}\big(\sum_{\substack{\boldsymbol{\sigma},\boldsymbol{\tau}\in\mathscr{O}_n \\ e_{\boldsymbol{\sigma}} e_{\boldsymbol{\tau}} = \pm e_{\boldsymbol{\beta}}}} \mathrm{sgn}(\boldsymbol{\sigma}\boldsymbol{\tau}) a_{\boldsymbol{\sigma}} f_{\boldsymbol{\tau}}\big) + S_{\boldsymbol{\alpha}}\big(\langle g\rangle_{\boldsymbol{\beta}}\big)\Big) \\ &= \sum_{\boldsymbol{\alpha},\boldsymbol{\sigma},\boldsymbol{\tau}\in\mathscr{O}_n} e_{\boldsymbol{\sigma}} e_{\boldsymbol{\tau}} e_{\boldsymbol{\alpha}} a_{\sigma} S_{\boldsymbol{\alpha}}(f_{\tau}) + \sum_{\boldsymbol{\alpha},\boldsymbol{\beta}\in\mathscr{O}_n} e_{\boldsymbol{\beta}} e_{\boldsymbol{\alpha}} S_{\boldsymbol{\alpha}}(g_{\boldsymbol{\beta}}) \\ &= aS(f) + S(g), \end{aligned}$$

where $e_{\boldsymbol{\beta}} = \mathrm{sgn}(\boldsymbol{\sigma}\boldsymbol{\tau}) e_{\boldsymbol{\sigma}} e_{\boldsymbol{\tau}}$. Hence S is a left-linear operator on the Hilbert Clifford module.

Let us now assume that S is a left-linear operator on the Hilbert Clifford module. For an element $a = \sum_{\boldsymbol{\alpha}\in\mathscr{O}_n} e_{\boldsymbol{\alpha}} a_{\boldsymbol{\alpha}}$ of the Clifford module denote

$$\underline{a}_{\boldsymbol{\beta}} := -\sum_{\substack{\boldsymbol{\alpha}\in\mathscr{O}_n \\ \boldsymbol{\alpha}\neq\boldsymbol{\beta}}} e_{\boldsymbol{\alpha}} a_{\boldsymbol{\alpha}} + e_{\boldsymbol{\beta}} a_{\boldsymbol{\beta}}.$$

Then $a + \underline{a}_{\boldsymbol{\alpha}} = 2e_{\boldsymbol{\alpha}}\langle a\rangle_{\boldsymbol{\alpha}}$ and consequently $a = \frac{1}{2}\sum_{\boldsymbol{\alpha}\in\mathscr{O}_n} a + \underline{a}_{\boldsymbol{\alpha}}$. Furthermore, since S is left-linear

$$\begin{aligned} S(f) &= S\Big(\frac{1}{2}\sum_{\boldsymbol{\alpha}\in\mathscr{O}_n} f + \underline{f}_{\boldsymbol{\alpha}}\Big) = \frac{1}{2}\sum_{\boldsymbol{\alpha}\in\mathscr{O}_n} S\big(f + \underline{f}_{\boldsymbol{\alpha}}\big) = \frac{1}{4}\sum_{\boldsymbol{\alpha},\boldsymbol{\beta}\in\mathscr{O}_n} \big(S + \underline{S}_{\boldsymbol{\alpha}}\big)\big(f + \underline{f}_{\boldsymbol{\beta}}\big) \\ &= \frac{1}{4}\sum_{\boldsymbol{\alpha},\boldsymbol{\beta}\in\mathscr{O}_n} e_{\boldsymbol{\alpha}}\Big\langle S\big(e_{\boldsymbol{\beta}}\langle f\rangle_{\boldsymbol{\beta}}\big)\Big\rangle_{\boldsymbol{\alpha}} = \sum_{\boldsymbol{\alpha},\boldsymbol{\beta}\in\mathscr{O}_n} e_{\boldsymbol{\alpha}}\Big\langle e_{\boldsymbol{\beta}} S\big(e_0\langle f\rangle_{\boldsymbol{\beta}}\big)\Big\rangle_{\boldsymbol{\alpha}} \\ &= \sum_{\boldsymbol{\alpha},\boldsymbol{\beta}\in\mathscr{O}_n} \mathrm{sgn}(\boldsymbol{\beta}\boldsymbol{\gamma}) e_{\boldsymbol{\alpha}} \langle S\big(e_0\langle f\rangle_{\boldsymbol{\beta}}\big)\rangle_{\boldsymbol{\gamma}}, \end{aligned}$$

where $\boldsymbol{\gamma} \in \mathscr{O}_n : \mathrm{sgn}(\boldsymbol{\beta}\boldsymbol{\gamma}) e_{\boldsymbol{\beta}} e_{\boldsymbol{\gamma}} = e_{\boldsymbol{\alpha}}$ and $\mathrm{sgn}(\boldsymbol{\beta}\boldsymbol{\gamma})$ is the sign such that $\mathrm{sgn}(\boldsymbol{\beta}\boldsymbol{\gamma}) e_{\boldsymbol{\beta}} e_{\boldsymbol{\gamma}} = e_{\boldsymbol{\alpha}}$ for some $\boldsymbol{\alpha} \in \mathscr{O}_n$.

Finally for $f \in H$ and $\boldsymbol{\alpha} \in \mathscr{O}_n$ denote $S_{\boldsymbol{\alpha}}(f) := \langle S(e_0 f)\rangle_{\boldsymbol{\alpha}}$.

Suppose now $\forall \boldsymbol{\alpha} \in \mathscr{O}_n \exists C_{\boldsymbol{\alpha}} \in \mathbb{R}^+ : \|S_{\boldsymbol{\alpha}} g\|^2 \leq C_{\boldsymbol{\alpha}} \|g\|^2,\ \forall g \in H$. Then for all $f \in H_n$

$$\begin{aligned}\|Sf\|^2 &= \sum_{\boldsymbol{\alpha}\in\mathscr{O}_n} \left\|\langle S(f)\rangle_{\boldsymbol{\alpha}}\right\|^2 = \sum_{\boldsymbol{\alpha}\in\mathscr{O}_n} \left\|\sum_{\tau\sigma=\boldsymbol{\alpha}} S_\sigma(f_\tau)\right\|^2 \\ &\leq \sum_{\boldsymbol{\alpha}\in\mathscr{O}_n}\sum_{\tau\sigma=\boldsymbol{\alpha}} \left\|S_\sigma(f_\tau)\right\|^2 \leq \sum_{\boldsymbol{\alpha}\in\mathscr{O}_n}\sum_{\tau\sigma=\boldsymbol{\alpha}} C_\sigma \left\|f_\tau\right\|^2 \leq C\sum_{\boldsymbol{\tau}} \|f_\tau\|^2 \\ &= C\|f\|^2,\end{aligned}$$

for some $C \in \mathbb{R}^+$.

On the other hand assume $\exists \boldsymbol{\alpha} \in \mathscr{O}_n$, $\{g_l\}_{l\in\mathbb{N}} \subset H$, $\|g_l\| = 1$ such that $\{\|S_{\boldsymbol{\alpha}}(g_l)\|\}_{l\in\mathbb{N}}$ is unbouded.

Let $f_l := g_l e_0$. Then

$$\|Sf_l\| = \sum_{\boldsymbol{\beta}\in\mathscr{O}_n} \|S_{\boldsymbol{\beta}}(e_0 g_l)\|^2 \geq \|S_\alpha(g)\|^2,$$

whence the series $\{\|Sf_l\|\}_{l\in\mathbb{N}}$ is unbounded. Since $\|S\|$ would be an upper bound for this series, S is unbounded.

□

Corollary 3.2.8 (The adjoint operator)
LET $T \in L(H_n, G_n)$.

THEN *its adjoint* T^* *with respect to the scalar product* $(\cdot,\cdot)$ *is equal to its adjoint with respect to the form* $\langle\!\langle \cdot,\cdot \rangle\!\rangle$.

$$(Tx, y) = (x, T^*y),\ \forall x, y \in H_n \quad \Leftrightarrow \quad \langle\!\langle Tx, y\rangle\!\rangle = \langle\!\langle x, T^*y\rangle\!\rangle,\ \forall x, y \in H_n.$$

Proof. Let $T \in L(H_n, G_n)$ and fix $y \in G_n$. Then $h(x) := \langle\!\langle Tx, y\rangle\!\rangle$ is a Clifford functional. As a consequence $l(x) = \langle\!\langle h(x)\rangle\!\rangle_0$ is a $\mathbb{C}$-linear functional. Now $l(x) = (Tx, y) = (x, T^*y)$. Furthermore $g(x) := \langle\!\langle x, T^*y\rangle\!\rangle$ is a Clifford functional and $\langle g(x)\rangle_0 = (x, T^*y) = l(x)$. By Proposition 3.2.2(ii) we know that $g(x) = \widetilde{l}(x) = h(x)$. □

Self-adjoint operators on Hilbert modules

Definition 3.2.9 (Unitary and self-adjoint operators)

- LET $T \in L(H_n, G_n)$.
  THEN T *is called* **unitary** *iff* $TT^* = \mathrm{Id}_{G_n}$, $T^*T = \mathrm{Id}_{H_n}$.
- LET $T \in L(H_n) := L(H_n, H_n)$.
  THEN T *is called* **self-adjoint** *iff* $T = T^*$.

Proposition 3.2.10 (Characterization of self-adjoint operators)
LET $T \in L(H_n)$.

THEN *the following are equal.*

(i) T *is self-adjoint.*

(ii) $\langle\!\langle Tx, x\rangle\!\rangle = \overline{\langle\!\langle Tx, x\rangle\!\rangle}$, $\forall x \in H_n$.

(iii) *For any* $\boldsymbol{\alpha} : \overline{e_{\boldsymbol{\alpha}}} = e_{\boldsymbol{\alpha}}$ *the operator* $T_{\boldsymbol{\alpha}} \in L(H)$ *is self-adjoint. For any* $\boldsymbol{\alpha} : \overline{e_{\boldsymbol{\alpha}}} = -e_{\boldsymbol{\alpha}}$ *the operator* $iT_{\boldsymbol{\alpha}} \in L(H)$ *is self-adjoint.*

Proof. Let $x, y \in H_n$.

(i)⇒ (ii) $\langle\!\langle Tx, x\rangle\!\rangle = \langle\!\langle x, T^* x\rangle\!\rangle = \langle\!\langle x, Tx\rangle\!\rangle = \overline{\langle\!\langle Tx, x\rangle\!\rangle}$.

(ii)⇒ (i) Let $a \in \mathbb{C}_n$.

$$\begin{aligned} &\langle\!\langle T(x+ay), x+ay\rangle\!\rangle = \langle\!\langle Tx, x\rangle\!\rangle + \langle\!\langle Tx, y\rangle\!\rangle\overline{a} + a\langle\!\langle Ty, x\rangle\!\rangle + a\langle\!\langle Ty, y\rangle\!\rangle\overline{a} \\ \overset{(ii)}{\Longrightarrow}\; &\overline{\langle\!\langle T(x+ay), x+ay\rangle\!\rangle} = \langle\!\langle Tx, x\rangle\!\rangle + a\langle\!\langle y, Tx\rangle\!\rangle + \langle\!\langle x, Ty\rangle\!\rangle\overline{a} + a\langle\!\langle Ty, y\rangle\!\rangle\overline{a}. \end{aligned}$$

Applying (ii) yields

$$\langle\!\langle Tx, y\rangle\!\rangle\overline{a} + a\langle\!\langle Ty, x\rangle\!\rangle = a\langle\!\langle y, Tx\rangle\!\rangle + \langle\!\langle x, Ty\rangle\!\rangle\overline{a}.$$

Choosing $a = 1$ we see that

$$\langle\!\langle Tx, y\rangle\!\rangle + \langle\!\langle Ty, x\rangle\!\rangle = \langle\!\langle y, Tx\rangle\!\rangle + \langle\!\langle Tx, y\rangle\!\rangle. \tag{3.2}$$

Now we use the fact that H_n is a sub-module of every Clifford-Hilbert module H_ν, such that $n \le \nu \in \mathbb{N}$. We choose $\nu \in \mathbb{N}$ such that ν and $\frac{\nu(\nu+1)}{2}$ are odd. Then $e_\nu := e_{1\ldots\nu}$ is an element of the commutator of the Clifford algebra $\mathbb{C}_\nu$, and $\overline{e_\nu} = -e_\nu$. Choosing $a = e_\nu$ we have

$$e_\nu\langle\!\langle y, Tx\rangle\!\rangle - e_\nu\langle\!\langle x, Ty\rangle\!\rangle = e_\nu\langle\!\langle Ty, x\rangle\!\rangle - e_\nu\langle\!\langle Tx, y\rangle\!\rangle. \tag{3.3}$$

Now by multiplying (3.3) by $\overline{e_\nu}$ from the left, and adding (3.2), (i) follows.

(iii)⇒ (ii)

$$\begin{aligned} \langle\!\langle Tx, x\rangle\!\rangle &= \sum_{\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\gamma}\in\mathcal{O}_n} e_\alpha e_{\boldsymbol{\beta}}\overline{e_{\boldsymbol{\gamma}}}\langle T_{\boldsymbol{\beta}} x_{\boldsymbol{\alpha}}, x_{\boldsymbol{\gamma}}\rangle \\ &\overset{(iii)}{=} \sum_{\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\gamma}\in\mathcal{O}_n} e_\alpha \overline{e_{\boldsymbol{\beta}}}\,\overline{e_{\boldsymbol{\gamma}}}\langle x_{\boldsymbol{\alpha}}, T_{\boldsymbol{\beta}} x_{\boldsymbol{\gamma}}\rangle = \langle\!\langle x, Tx\rangle\!\rangle \\ &= \overline{\langle\!\langle Tx, x\rangle\!\rangle} \end{aligned}$$

(ii)⇒ (iii) By Theorem 3.2.7(viii) $\forall \boldsymbol{\alpha} \in \mathcal{O}_n \exists T_{\boldsymbol{\alpha}} \in L(H):\; T = \sum_{\boldsymbol{\alpha}\in\mathcal{O}_n} e_{\boldsymbol{\alpha}} T_{\boldsymbol{\alpha}}$.

$$\langle\!\langle Tx, x\rangle\!\rangle = \sum_{\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\gamma}\in\mathcal{O}_n} e_\alpha e_{\boldsymbol{\beta}}\overline{e_{\boldsymbol{\gamma}}}\langle T_{\boldsymbol{\beta}} x_{\boldsymbol{\alpha}}, x_{\boldsymbol{\gamma}}\rangle$$

comparing the coefficients with

$$\langle\!\langle Tx, x\rangle\!\rangle = \overline{\langle\!\langle Tx, x\rangle\!\rangle} = \sum_{\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\gamma}\in\mathcal{O}_n} e_\alpha \overline{e_{\boldsymbol{\beta}}}\,\overline{e_{\boldsymbol{\gamma}}}\langle x_{\boldsymbol{\alpha}}, T_{\boldsymbol{\beta}} x_{\boldsymbol{\gamma}}\rangle$$

yields (iii).

□

Remark 3.2.11
In the second part of the last proof one could just as well have used $a = i$. The chosen approach has the advantage that it works with non complex Clifford-Hilbert modules, i.e. Hilbert-Modules of $\mathbb{R}_n$ over a Hilbert space, too.

EXAMPLE 3.2.2 (Multiples of the identity):
We want to consider operators of the form $a\mathrm{Id} : H_n \to H_n,\ x \mapsto xa$, where $a \in \mathbb{C}_n$. Proposition 3.2.10 yields that these operators are self adjoint, iff

$$a_{\boldsymbol{\alpha}} \in \begin{cases} \mathbb{R} & \forall \boldsymbol{\alpha} : \overline{e_{\boldsymbol{\alpha}}} = e_{\boldsymbol{\alpha}}, \\ i\mathbb{R} & \forall \boldsymbol{\alpha} : \overline{e_{\boldsymbol{\alpha}}} = -e_{\boldsymbol{\alpha}}. \end{cases}$$

Theorem 3.2.12
LET $T \in L(H_n)$ *be self-adjoint.*
THEN

$$\|T\| \leq \sup_{\|x\| \leq 1} |\langle\!\langle Tx, x \rangle\!\rangle|.$$

Furthermore $\sup_{\|x\|\leq 1} |\langle\!\langle Tx, x\rangle\!\rangle| \leq 2^{n/2}\|T\|$.

Proof. Since $\big(H_n, (\,\cdot\,,\,\cdot\,)\big)$ is a complex Hilbert space

$$\|T\| = \sup_{\|x\|\leq 1} |(Tx, x)| \leq \sup_{\|x\|\leq 1} |\langle\!\langle Tx, x\rangle\!\rangle|.$$

(See for example [56] V.5.7.)
Using Theorem 3.2.4(iii) we know that

$$\sup_{\|x\|\leq 1} |\langle\!\langle Tx, x\rangle\!\rangle| \leq 2^{n/2} \sup_{\|x\|\leq 1} \|Tx\|\,\|x\| = \|T\|.$$

□

Corollary 3.2.13
LET $T \in L(H_n)$ *be self-adjoint.*
THEN

$$\langle\!\langle Tx, x\rangle\!\rangle = 0 \quad \Leftrightarrow \quad T = 0.$$

3.2.4 The Fourier transform on $L^p(\mathbb{R}^n)_n$

We extend the Fourier transform in a standard way to Clifford algebra valued functions.

Definition 3.2.14 (Fourier transform on Clifford-Hilbert modules)
LET $f = \sum_{j\in\mathcal{O}_n} e_j f_j \in L^1(\mathbb{R}^n)_n$.
THEN *the* **Fourier transform** *of f is defined as*

$$\mathscr{F}(f)(\xi) = \widehat{f}(\xi) := \sum_{j\in\mathcal{O}_n} e_j \widehat{f}_j(\xi) = \int_{\mathbb{R}_n} f(x) e^{2\pi i\langle x,\xi\rangle}\,dx.$$

This definition is then extended to the corresponding L^p-modules, $1 \leq p < \infty$, in the usual way.

The Euclidean scalar product on $L^2(\mathbb{R}^n)_n$ is given by

$$(f,g) := \sum_{i\in\mathcal{O}_n} \langle f_i, g_i\rangle_{L^2(\mathbb{R}^n)} = \Re\Big(\int_{\mathbb{R}^n} f(x)\overline{g(x)}dx\Big).$$

This scalar product satisfies $\|f\|_2^2 = (f,f)$.

Theorem 3.2.15 (The Plancherel equality)
LET $f,g \in L^2(\mathbb{R}^n)_n$.
THEN $\langle\langle f,g\rangle\rangle = \langle\langle \hat{f},\hat{g}\rangle\rangle$. *Consequently* $(f,g) = (\hat{f},\hat{g})$.

Proof. Let $f = \sum_{j\in\mathcal{O}_n} e_j f_j$, where $f_j \in L^2(\mathbb{R}^n,\mathbb{C}) \quad \forall j \in \mathbb{N}_n$. Then

$$\langle\langle f,g\rangle\rangle = \sum_{j,k\in\mathcal{O}_n} e_j\overline{e_k}\langle f_j,g_k\rangle_{L^2(\mathbb{R}^n)} \overset{(i)}{=} \sum_{j,k\in\mathcal{O}_n} e_j\overline{e_k}\langle \hat{f_j},\hat{g_k}\rangle = \langle\langle \hat{f},\hat{g}\rangle\rangle.$$

In (i) we used the Plancherel inequality for $L^2(\mathbb{R}^n,\mathbb{C})$. □

3.3 Clifford analysis

Our goal is to define an analytical signal in $\mathbb{R}^n$. In Theorem 2.1.2 and Theorem 2.1.7 we have seen that the Riesz transforms are a suitable extension of the Hilbert transform to higher dimensions. In this chapter we will introduce a hyper-complex analysis that is built on a set of generalized Cauchy Riemann equations that can characterize the analytical signal in one and higher dimensions as a boundary value of a function satisfying these equations.

This hyper-complex analysis is called Clifford analysis. We will only give a short introduction to Clifford analysis. For details see [13], [21] and [19].

For $n \in \mathbb{N}$ we will consider the Clifford algebras $\mathbb{R}_n$, and functions in $C^1(\mathbb{R}^{n+1},\mathbb{R}^{n+1})$, where $\mathbb{R}^{n+1}$ is identified with the real vector space of paravectors $\text{span}\{e_\alpha\}_{\alpha=0}^n$ in $\mathbb{R}_n$. In this manner, we will not encounter any zero divisors.

Functions $f : \mathbb{R}^{n+1} \mapsto \mathbb{R}^{n+1}$ in this setting are of the form

$$f(x) = \sum_{\alpha=0}^{n} f_\alpha(x)e_\alpha,$$

where $f_\alpha : \mathbb{R}^{n+1} \to \mathbb{R}$, $\alpha = 0,\dots,n$.

The **Dirac operator** is given by

$$D := \sum_{\beta=1}^{n} \frac{\partial}{\partial x_\beta} e_\beta$$

and the **Cauchy-Riemann operator** by

$$\partial := \sum_{\beta=0}^{n} \frac{\partial}{\partial x_\beta} e_\beta,$$

clearly $\partial = \frac{\partial}{\partial x_0} + D$.

Since multiplication in this algebra is not commutative, these operators may act from the left or from the right.

We will assume that a differential operator defined as $d := \sum_{\beta=0}^{n} \frac{\partial}{\partial x_\beta} e_\beta$ usually **acts from the left** by

$$df(x) = \sum_{\boldsymbol{\alpha}\in\mathcal{O}_n} \sum_{\beta} \frac{\partial f_{\boldsymbol{\alpha}}}{\partial x_\beta}(x) e_\beta e_{\boldsymbol{\alpha}}.$$

The operator **acting** on a function **from the right** is then denoted by

$$(fd)(x) = \sum_{\boldsymbol{\alpha}\in\mathcal{O}_n} \sum_{\beta=0}^{n} \frac{\partial f_{\boldsymbol{\alpha}}}{\partial x_\beta}(x) e_{\boldsymbol{\alpha}} e_\beta.$$

Definition 3.3.1 (Monogenic Function)
LET $f : \mathbb{R}^{n+1} \to \mathbb{R}_n$ *be a partially differentiable function such that f fulfills the Cauchy-Riemann type equation*

$$\partial f(x) = \sum_{\boldsymbol{\alpha}\in\mathcal{O}_n} \sum_{\beta=0}^{n} \frac{\partial f_{\boldsymbol{\alpha}}}{\partial x_\beta} e_\beta e_{\boldsymbol{\alpha}}(x) = 0,$$

where $f(x) = \sum_{\alpha\in\mathcal{O}_n} e_\alpha f_\alpha(x)$.

THEN f *is called (left-)* **monogenic**.

A partially differentiable function $f : \mathbb{R}^{n+1} \to \mathbb{R}_n$ *fulfilling the Cauchy-Riemann-type equation* $(f\partial) = 0$ *is called* **right monogenic**.

Theorem 3.3.2 (Left- and right monogenic functions)
LET $f : \mathbb{R}^{n+1} \to \mathbb{R}^{n+1}$ *be a left monogenic function.*

THEN f *is right monogenic.*

Proof. Let f be a left monogenic function. Then

$$\partial f = \sum_{\alpha,\beta=0}^{n} e_\alpha e_\beta \frac{\partial}{\partial\alpha} f_\beta = \sum_{\alpha=0}^{n} e_\alpha e_\alpha \frac{\partial}{\partial\alpha} f_\alpha + \sum_{\alpha\in\{1,\ldots,n\}} e_\alpha \Big(\frac{\partial}{\partial\alpha} f_0 + \frac{\partial}{\partial 0} f_\alpha\Big) + \sum_{\substack{\alpha,\beta\in\{1,\ldots,n\}\\ \alpha<\beta}} e_\alpha e_\beta \Big(\frac{\partial}{\partial\alpha} f_\beta - \frac{\partial}{\partial\beta} f_\alpha\Big)$$
$$= 0.$$

It follows that

$$\frac{\partial}{\partial 0} f_0 - \sum_{\alpha\in\{1,\ldots,n\}} \frac{\partial}{\partial\alpha} f_\alpha = 0,$$

$$\Big(\frac{\partial}{\partial\alpha} f_0 + \frac{\partial}{\partial 0} f_\alpha\Big) = 0, \ \forall \alpha \in \{1,\ldots,n\}$$

and

$$\frac{\partial}{\partial\alpha} f_\beta = -\frac{\partial}{\partial\beta} f_\alpha, \ \forall \alpha \neq \beta \in \{1,\ldots,n\}.$$

As a consequence

$$0 = \sum_{\alpha=0}^{n} e_\alpha e_\alpha \frac{\partial}{\partial\alpha} f_\alpha + \sum_{\alpha\in\{1,\ldots,n\}} e_\alpha \Big(\frac{\partial}{\partial\alpha} f_0 + \frac{\partial}{\partial 0} f_\alpha\Big) + \sum_{\substack{\alpha,\beta\in\{1,\ldots,n\}\\ \alpha<\beta}} e_\beta e_\alpha \Big(\frac{\partial}{\partial\alpha} f_\beta - \frac{\partial}{\partial\beta} f_\alpha\Big) = \sum_{\alpha,\beta=0}^{n} e_\alpha e_\beta \frac{\partial}{\partial\alpha} f_\beta$$
$$= (f\partial).$$

□

Remark 3.3.3
As a consequence of Theorem 3.3.2 we will henceforth simply speak of monogenic functions. Theorem 3.3.2 does not state that $\partial f = (f\partial)$

Corollary 3.3.4
Let $f \in C^1(\mathbb{R}^{n+1}, \mathbb{R}^{n+1})$*. Then monogenicity is equivalent to a set of Cauchy-Riemann equations for the component functionals* f_α *:*

$$\begin{aligned} \frac{\partial f_{\boldsymbol{\alpha}}}{\partial x_\beta}(x) &= \frac{\partial f_\beta}{\partial x_\alpha}(x),\ \forall \alpha, \beta = 1, \ldots, n \\ \frac{\partial f_\alpha}{\partial x_0}(x) + \frac{\partial f_0}{\partial x_\alpha}(x) &= 0,\ \forall \alpha = 1, \ldots, n \\ \sum_{\alpha=1}^{n} \frac{\partial f_\alpha}{\partial x_\alpha}(x) &= \frac{\partial f_0}{\partial x_0}(x). \end{aligned} \tag{3.4}$$

Proof. We have proven this in the proof of Theorem 3.3.2. □

Remark 3.3.5
Notice that the Dirac operator $D = \sum_{\beta=1}^{n} \frac{\partial}{\partial x_\beta} e_\beta$ *yields the Cauchy-Riemann like system* (2.2)*. That is let* $f = \sum_{\alpha=1}^{n+1} e_\alpha f_\alpha \in L^2(\mathbb{R}^n, \mathbb{R}^{n+1}) \subset L^2(\mathbb{R}^n)_{n+1}$ *be continuously differentiable. Then*

$$Df = 0$$

iff

$$\begin{aligned} \sum_{\alpha=1}^{n+1} \frac{\partial f_\alpha}{\partial x_\alpha} &= 0, \\ \frac{\partial f_\alpha}{\partial x_\beta} &= \frac{\partial f_\beta}{\partial x_\alpha};\ \alpha, \beta = 1, \ldots, n+1. \end{aligned}$$

This is exactly (2.2) *if we set* $u_0 = f_{n+1}$ *and* $u_\alpha = f_\alpha$, $\forall \alpha = 1, \ldots, n$.

EXAMPLE 3.3.1 (Monomials):
Let $1 < n \in \mathbb{N}$. Then the monomials x^d, $x \in \mathbb{R}^{n+1}$; $d \in \mathbb{N}$ are not monogenic on the whole space:
Let $d = 1$, that is $x^1 = \sum_{\alpha=0}^{n} e_\alpha x_\alpha$. Then

$$\partial x = \frac{\partial x_0}{\partial x_0} + \sum_{\alpha=1}^{n} e_\alpha e_\alpha \frac{\partial x_\alpha}{\partial x_\alpha} = 1 - n.$$

Let $d = 2$, that is

$$\begin{aligned} x^2 &= (\sum_{\alpha=0}^{n} e_\alpha x_\alpha)(\sum_{\beta=0}^{n} e_\beta x_\beta) \\ &= \sum_{\alpha=0}^{n} e_\alpha^2 x_\alpha^2 + \sum_{\alpha,\beta=0}^{n} (e_\alpha e_\beta + e_\beta e_\alpha) x_\alpha x_\beta \\ &= x_0^2 - \sum_{\alpha=1}^{n} x_\alpha^2 + 2\sum_{\alpha=1}^{n} e_\alpha x_\alpha x_0 + \sum_{\alpha,\beta=1}^{n} (e_\alpha e_\beta - e_\alpha e_\beta) x_\alpha x_\beta \\ &= x_0^2 - \sum_{\alpha=1}^{n} x_\alpha^2 + 2\sum_{\alpha=1}^{n} e_\alpha x_\alpha x_0. \end{aligned}$$

Then

$$\begin{aligned}\partial x^2 &= \frac{\partial(x_0^2 + 2\sum_{\alpha=1}^n e_\alpha x_\alpha x_0)}{\partial x_0} + \sum_{\alpha=1}^n e_\alpha \frac{\partial(2e_\alpha x_\alpha x_0 - x_\alpha^2)}{\partial x_\alpha} \\ &= 2x_0 + 2\sum_{\alpha=1}^n e_\alpha x_\alpha - 2\sum_{\alpha=1}^n e_\alpha x_\alpha - 2\sum_{\alpha=1}^n x_0 \\ &= 2x_0(1-n).\end{aligned}$$

For the complex case, i.e. $n = 1$ and $\mathbb{R}_n = \mathbb{C}$, where $1 - n = 0$, the terms containing $(1 - n)$ vanish, and thus the monomials are analytical, as is well known.

EXAMPLE 3.3.2 (Fueter variables):
A **Fueter variable** is defined as $z_\alpha(x) := x_\alpha - e_\alpha x_0$. This function is clearly monogenic and, if we define $z_0 = x_0$ we gain a basis for $\mathbb{R}^{n+1}$ consisting of monogenic elements.

The product of two monogenic functions is not in general everywhere monogenic on $\mathbb{R}^{n+1}$. Take for example the product of the Fueter variables

$$z_\alpha z_\beta(x) = x_\alpha x_\beta - e_\beta x_0 x_\alpha - e_\alpha x_0 x_\beta + e_{\alpha\beta} x_0^2.$$

Then,

$$\begin{aligned}\partial z_\alpha z_\beta(x) &= \frac{\partial(e_{\alpha\beta}x_0^2 - e_\beta x_0 x_\alpha - e_\alpha x_0 x_\beta)}{\partial x_0} + e_\alpha \frac{\partial(x_\alpha x_\beta - e_\beta x_0 x_\alpha)}{\partial x_\alpha} + e_\beta \frac{\partial(x_\alpha x_\beta - e_\alpha x_0 x_\beta)}{\partial x_\beta} \\ &= 2e_{\alpha\beta}x_0 - e_\beta x_\alpha - e_\alpha x_\beta + e_\alpha x_\beta - e_{\alpha\beta}x_0 + e_\beta x_\alpha + e_{\alpha\beta}x_0 \\ &= 2e_{\alpha\beta}x_0.\end{aligned}$$

3.3.1 The exponential function

Theorem 3.3.6 (Paravector-valued Exponential Function)
By

$$\exp : \mathbb{R}^{n+1} \to \mathbb{R}^{n+1},\ x \mapsto \sum_{k\in\mathbb{N}_0} \frac{x^k}{k!}$$

an **exponential function for paravectors** *which is of the form*

$$\exp(x) = e^{x_0}\left(\cos|\vec{x}| + \frac{\vec{x}}{|\vec{x}|}\sin|\vec{x}|\right),\ \forall x \in \mathbb{R}^{n+1} \tag{3.5}$$

is defined.

This function is not monogenic as follows from Example 3.3.1 but has the usual properties of an exponential function if restricted to the subspace $\mathbb{R} + \vec{x}\mathbb{R}$, which can be identified with $\mathbb{C}$ since $\frac{\vec{x}}{|\vec{x}|} \in S^n$. It holds true that

(i) $\exp(x+y) = \exp(x)\exp(y),\ \forall xy = yx,\ x, y \in \mathbb{R}^{n+1}$;

(ii) $\exp(-x)\exp(x) = 1,\ \exp(x) \neq 0$;

(iii) $\exp(kx) = (\exp(x))^k$;

(iv) $\partial \exp(x) = e^{x_0}\frac{n-1}{|\vec{x}|}\sin(|\vec{x}|)$.

Proof. Let $x = x_0 + \vec{x} \in \mathbb{R}^{n+1}$ and $i := \frac{\vec{x}}{|\vec{x}|}$. It follows that $i^2 = -1$. Note that the elements $\frac{x^k}{k!}$ of the series defining the exponential function are contained in the subalgebra $\mathbb{R} + i\mathbb{R} \subset \mathbb{R}_n$, which is isomorphic to $\mathbb{C}$.

Hence writing $x = x_0 + i|\vec{x}|$ we know that

$$\exp(x) = e^{x_0}\big(\cos(|\vec{x}|) + i\sin(|\vec{x}|)\big) = e^{x_0}\Big(\cos(|\vec{x}|) + \frac{\vec{x}}{|\vec{x}|}\sin(|\vec{x}|)\Big).$$

Now equations (ii) and (iii) follow from the corresponding properties of the exponential function in $\mathbb{C}$.

(i) is easy to compute, the proof is exactly the same as in the case of the exponential function of matrices.

(iv)

$$\begin{aligned}
\partial\exp(x) &= \sum_{\alpha=0}^{n} e_\alpha \frac{\partial}{\partial x_\alpha}\Big(e^{x_0}\big(\cos(|\vec{x}|) + \sum_{\beta=1}^{n} e_\beta \frac{x_\beta}{|\vec{x}|}\sin(|\vec{x}|)\big)\Big)\\
&= \frac{\partial e^{x_0}\exp(\vec{x})}{\partial x_0} + e^{x_0}\Big(\sum_{\alpha=1}^{n} e_\alpha \frac{\partial\cos(|\vec{x}|)}{\partial x_\alpha} + \sum_{\beta=1}^{n} e_\alpha e_\beta \frac{\partial}{\partial x_\alpha}\frac{x_\beta}{|\vec{x}|}\sin(|\vec{x}|)\Big)\\
&= e^{x_0}\big(\cos(|\vec{x}|) + \frac{\vec{x}}{|\vec{x}|}\sin(|\vec{x}|)\big) + e^{x_0}\Big(-\sum_{\alpha=1}^{n} e_\alpha \frac{x_\alpha\sin(|\vec{x}|)}{|\vec{x}|} + \sum_{\beta=1}^{n} e_\alpha e_\beta \frac{x_\beta}{|\vec{x}|}\cos(|\vec{x}|)\frac{x_\alpha}{|\vec{x}|}\\
&\quad + \sin(|\vec{x}|)\big(\sum_{\beta\neq\alpha} e_\alpha e_\beta \frac{-x_\beta x_\alpha}{|\vec{x}|^3} + \frac{x_\alpha^2}{|\vec{x}|^3}\big)\Big)\\
&= e^{x_0}\Big(\cos(|\vec{x}|) + \frac{\vec{x}}{|\vec{x}|}\sin(|\vec{x}|) - \sum_{\alpha=1}^{n} e_\alpha \frac{x_\alpha\sin(|\vec{x}|)}{|\vec{x}|} - \frac{x_\alpha}{|\vec{x}|}\cos(|\vec{x}|)\frac{x_\alpha}{|\vec{x}|} + \sin(|\vec{x}|)\frac{x_\alpha^2}{|\vec{x}|^3}\Big)\\
&= e^{x_0}\frac{n-1}{|\vec{x}|}\sin(|\vec{x}|).
\end{aligned}$$

This holds true, since $\frac{\partial}{\partial x_\alpha}\frac{x_\alpha}{|\vec{x}|} = (\frac{1}{|\vec{x}|} - \frac{x_\alpha^2}{|\vec{x}|^3}) = \sum_{\beta\neq\alpha}\frac{x_\beta^2}{|\vec{x}|^3}$ and since the terms containing $e_\alpha e_\beta$ are cancelled by the terms containing $e_\beta e_\alpha = -e_\alpha e_\beta$. □

Definition 3.3.7 (Phase of a Paravector)
LET $x = x_0 + \vec{x} \in \mathbb{R}^{n+1}$ *be a paravector.*

THEN *a* **phase** ϕ, *a* **phase direction** $d \in \mathbb{R}_n$ *and an* **amplitude** $a \in \mathbb{R}^+$ *are defined by* $x = a(\cos(\phi) + d\sin(\phi))$, *whence* $a = |x|$, $d = \frac{\vec{x}}{|\vec{x}|}$ *and* $\phi = \arctan\big(\frac{x_0}{|\vec{x}|}\big)$. *(See Figure 3.1 for an illustration.)*

Remark 3.3.8 (Relation to spherical coordinates)
In $\mathbb{C}$ the phase is closely related to the polar coordinates, since $d = \operatorname{sgn}(\Im(x))$ we have a decomposition into an amplitude, which corresponds to the radius, the phase angle and a phase sign. Phase angle and phase sign together correspond to the angle in polar coordinates: $d\phi$ is equal to the angle in polar coordinates. This leads to a continuous phase angle ϕ. The complex conjugate $\overline{x}$ of a $x \in \mathbb{C}$ has the same phase angle ϕ as x, but the negative phase sign d.

For $x \in \mathbb{R}^{n+1}$, $x_n := x = a(cos(\phi_n) + d_n\sin(\phi_n))$ we can map the vector valued direction d^n to a paravector via $\mathbb{R}^n \mapsto \mathbb{R}^{(n-1)+1}$, $d_n = \sum_{i=1}^n e_i d_{n,i} \mapsto x_{n-1} = \sum_{i=1}^n e_{i-1} d_{n,i}$ and use the decomposition, where $a = 1$. This leads to a decomposition into one amplitude, n phase angels with range $[0,\pi[$ and one sign $d = d_1$. The phase angels $\phi_2,\ldots,\phi_n$ are equal to the angles of

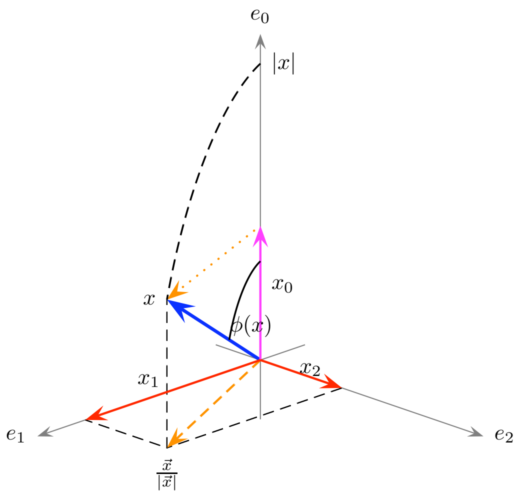


Figure 3.1: Phase ϕ and phase direction $\frac{\vec{x}}{\|\vec{x}\|}$ (direction of the dashed vector) of the 3-D vector $x = (x_0, x_1, x_2)$.

the spherical coordinates that range between $[0, \pi[$. $d_1\phi_1$ *is then equal to the spherical angle ranging between* $[0, 2\pi[$.
This yields a set of spherical coordinates for $\mathbb{R}^{n+1}$ *of the form* $x = (x_0, x_1, \ldots, x_n)$*, where*

$$\begin{aligned} x_0 &= a\cos(\phi_n) \\ x_1 &= a\sin(\phi_n)\cos(\phi_{n-1}) \\ x_2 &= \sin(\phi_n)\sin(\phi_{n-1})\cos(\phi_{n-2}) \\ &\vdots \\ x_{n-1} &= ad_1\sin(\phi_n)\cdots\sin(\phi_2)\cos(\phi_1) \\ x_n &= ad_1\sin(\phi_n)\cdots\sin(\phi_2)\sin(\phi_1). \end{aligned}$$

Explicitly in the case that $n+1=3$ *this yields*

$$x = a\big(\cos(\phi), d_\theta\sin(\phi)\cos(\theta), d_\theta\sin(\phi)\sin(\theta)\big).$$

This differs slightly from the convention we use in $\mathbb{R}^3$*. (See subsection A.1.2.)*

Chapter 4

Monogenic signals

The agenda of this chapter is to formulate monogenic signals as an extension of analytical signals to higher dimensions and to examine the properties of monogenic signals.

The first step is the definition of a hypercomplex Riesz transform R, given in Definition 4.1.1, that extends the Hilbert transform $i\mathscr{H}$. What is novel about this definition is that we define the Riesz transform as a left linear-operator on a Clifford-Hilbert module. As a consequence the (hypercomplex) Riesz transform we define is unitary and self-adjoint as is shown in Theorem 4.1.4. It follows that the Riesz transform is its own inverse. Furthermore, the Riesz transform is closely connected to differential operators as seen in a new result presented in Theorem 4.2.1.

We define monogenic signals in Definition 4.3.1, and in Definition 4.4.1 we state the decomposition of a monogenic signal into phase, amplitude, and phase direction. From the phase we derive in Definition 4.4.2 an instantaneous frequency for the monogenic signal. From the properties of the Riesz transforms it follows that a monogenic signal is generated by a projection $\mathscr{M} : f \mapsto f_m$ (Theorem 4.3.3) that is closely connected to Clifford analysis. We state these connections in section 4.5.

The final section will show how a monogenic signal is uniquely defined as the extension of an analytical signal by the condition that it is compatible with rotations, which generate the symmetry group we have to deal with in image analysis.

Some of the results of this chapter have been published in the peer reviewed article [24] and in [23].

4.1 Hypercomplex Riesz transforms

Depending on the kind of dilation we use, we need to define analytical signals in different ways. The key point here is to look at the invariances of the corresponding analog of the Hilbert transform.

The Riesz transform commutes with isotropic dilation and behaves like a vector under rotation. Thus, it is a tool of choice for a setting where rotations are important such as, for example, in image analysis.

Since we wish to be able to work with images and hence rotations we will now use the Riesz transform defined earlier to define a hyper-complex analytical signal. For this purpose, we introduce a hyper-complex Riesz transform, which plays a role similar to that of the Hilbert transform.

Definition 4.1.1 (Hypercomplex Riesz Transform)
LET $f \in L^p(\mathbb{R}^n, \mathbb{R}_n)$, $1 < p < \infty$.

THEN *we define the* **hypercomplex Riesz transform** *by*

$$\begin{aligned} R : & L^p(\mathbb{R}^n, \mathbb{R}_n) \to L^p(\mathbb{R}^n, \mathbb{R}_n) \\ & f \mapsto Rf = \sum_{\alpha=1}^{n} R_\alpha f e_\alpha = \sum_{\alpha=1}^{n} \sum_{\boldsymbol{\beta} \in \mathscr{O}_n} e_{\boldsymbol{\beta}} e_\alpha R_\alpha f_{\boldsymbol{\beta}}. \end{aligned}$$

Corollary 4.1.2 (Fourier multiplier of the hypercomplex Riesz transform)
LET $f \in L^p(\mathbb{R}^n, \mathbb{R}_n)$.

THEN *it follows from Theorem 2.1.2 (vii) that*

$$\mathscr{F}(Rf)(\xi) = \sum_{\alpha=1}^{n} \sum_{\boldsymbol{\beta} \in \mathscr{O}_n} e_{\boldsymbol{\beta}} e_\alpha \frac{i\xi_\alpha}{\|\xi\|} \widehat{f_{\boldsymbol{\beta}}}(\xi), \text{ f.a.a. } \xi \in \mathbb{R}^n.$$

I.e. the Fourier multiplier of the hypercomplex Riesz transform is

$$\widehat{R}(\xi) = \sum_{\alpha=1}^{n} e_\alpha \frac{i\xi_\alpha}{|\xi|}, \ \forall \xi = (\xi_1, \ldots, \xi_n) \in \mathbb{R}^n.$$

Corollary 4.1.3 (The hypercomplex Riesz transform of real valued functions)
LET f *be a real valued function* $f = e_0 f_0$, $f_0 \in L^2(\mathbb{R}^n, \mathbb{R})$.

THEN *the Riesz transform maps f into the space of 1-vectors.*

$$\begin{aligned} R : & L^p(\mathbb{R}^n, \mathbb{R}) \to L^p(\mathbb{R}^n, \mathbb{R}^n) \\ & f \mapsto Rf = \sum_{\alpha=1}^{n} e_\alpha R_\alpha f. \end{aligned}$$

We will now proof that the Riesz transform is a self-adjoint, unitary and hence bijective operator.

Theorem 4.1.4 (The hypercomplex Riesz transform is self-adjoint and unitary)
Let $f \in L^2(\mathbb{R}^n, \mathbb{C}_n)$. *Then*

(i) R is self-adjoint, i.e. $\langle\langle Rf, g \rangle\rangle = \langle\langle f, Rg \rangle\rangle$,

(ii) R is unitary, i.e. $\langle\langle Rf, Rg \rangle\rangle = \langle\langle f, g \rangle\rangle$ *and hence* $\|Rf\| = \|f\|$.

Proof. Let $f, g \in L^2(\mathbb{R}^n, \mathbb{R}_n)$. Then ad (i)

$$\begin{aligned}\langle\langle \widehat{Rf}, \widehat{g}\rangle\rangle &= \sum_{\alpha\in\{1,\ldots,n\}} \sum_{\boldsymbol{\beta},\boldsymbol{\gamma}\in\mathscr{O}_n} e_{\boldsymbol{\beta}} e_\alpha \overline{e_{\boldsymbol{\gamma}}} \int_{\mathbb{R}^n} \frac{ix_\alpha}{|x|} \widehat{f_{\boldsymbol{\beta}}}(x)\overline{\widehat{g_{\boldsymbol{\gamma}}}}(x)dx \\ &= \sum_{\alpha\in\{1,\ldots,n\}} \sum_{\boldsymbol{\beta},\boldsymbol{\gamma}\in\mathscr{O}_n} e_{\boldsymbol{\beta}} (-\overline{e_\alpha})\overline{e_{\boldsymbol{\gamma}}} \int_{\mathbb{R}^n} \widehat{f_{\boldsymbol{\beta}}}(x) \overline{\frac{-ix_\alpha}{|x|}\widehat{g_{\boldsymbol{\gamma}}}}(x)dx \\ &= \sum_{\alpha\in\{1,\ldots,n\}} \sum_{\boldsymbol{\beta},\boldsymbol{\gamma}\in\mathscr{O}_n} e_{\boldsymbol{\beta}} \overline{e_{\boldsymbol{\gamma}} e_\alpha} \int_{\mathbb{R}^n} \widehat{f_{\boldsymbol{\beta}}}(x) \overline{\frac{ix_\alpha}{|x|}\widehat{g_{\boldsymbol{\gamma}}}(x)}dx \\ &= \langle\langle \widehat{f}, \widehat{Rg}\rangle\rangle.\end{aligned}$$

ad (ii)

$$\begin{aligned}\langle\langle \widehat{Rf}, \widehat{Rg}\rangle\rangle &= \langle\langle \widehat{RRf}, \widehat{g}\rangle\rangle = \sum_{\beta,\gamma=1}^{n} \sum_{\boldsymbol{\alpha},\boldsymbol{\delta}\in\mathscr{O}_n} e_{\boldsymbol{\alpha}} e_\gamma e_\beta \overline{e_{\boldsymbol{\gamma}}} \int_{\mathbb{R}^n} \frac{-\xi_\beta\xi_\gamma}{\|\xi\|^2} \widehat{f_{\boldsymbol{\alpha}}}(\xi)\overline{\widehat{g_{\boldsymbol{\delta}}}}(\xi)d\xi \\ &= \sum_{\boldsymbol{\alpha},\boldsymbol{\delta}\in\mathscr{O}_n} e_{\boldsymbol{\alpha}} \Big(\sum_{1\le\beta<\gamma\le n} \underbrace{(e_\gamma e_\beta - e_\beta e_\gamma)}_{=0} \int_{\mathbb{R}^n} \frac{-\xi_\beta\xi_\gamma}{\|\xi\|^2} \widehat{f_{\boldsymbol{\alpha}}}(\xi)\overline{\widehat{g_{\boldsymbol{\delta}}}}(\xi)d\xi \\ &\quad + \sum_{\beta=1}^{n} e_{\boldsymbol{\alpha}} \underbrace{e_\beta^2}_{=-1} \overline{e_{\boldsymbol{\delta}}} \frac{-\xi_\beta^2}{\|\xi\|^2} \widehat{f_{\boldsymbol{\alpha}}}(\xi)\overline{\widehat{g_{\boldsymbol{\delta}}}}(\xi)\Big) \\ &= \langle\langle \widehat{f}, \widehat{g}\rangle\rangle.\end{aligned}$$

Now $\|Rf\|^2 = (Rf, Rf) = \Re\big(\langle\langle Rf, Rf\rangle\rangle\big) = \Re\big(\langle\langle f, f\rangle\rangle\big) = (f, f) = \|f\|^2$.

□

4.2 Differential operators and Riesz transforms

The Riesz transform is closely connected to the Dirac and to the Laplace operator. The following equalities hold only for the hypercomplex Riesz transform on Hilbert-Clifford modules.

Theorem 4.2.1
LET $f \in W^1(\mathbb{R}_n)$.
THEN

(i) $DRf = (-\triangle)^{1/2} f \in L^2(\mathbb{R}^n)$.

(ii) $\frac{\partial f}{\partial x_j} = R_j RDf \in L^2(\mathbb{R}^n)$.

Proof. ad(i) It is almost everywhere true that

$$\begin{aligned}\mathscr{F}(DRf)(\xi) &= \sum_{j=1}^{n} 2\pi e_j i\xi_j \sum_{k=1}^{n} \frac{ie_k\xi_k}{|\xi|}\widehat{f}(\xi) \\ &= 2\pi\Big(\sum_{j=1}^{n} \frac{\xi_j^2}{|\xi|} - \sum_{j<k}(e_{jk}+e_{kj})(\xi_j\xi_k)\Big)\widehat{f}(\xi) = 2\pi|\xi|\widehat{f}(\xi) \\ &= \mathscr{F}\big((-\triangle)^{1/2} f\big)(\xi).\end{aligned}$$

ad(ii) Again almost everywhere, we have that

$$\begin{aligned}\mathscr{F}(R_j RDf)(\xi) &= \frac{i\xi_j}{|\xi|}\sum_{k=1}^{n}\frac{e_k i\xi_k}{|\xi|}\sum_{l=1}^{n}(2\pi i)e_l\xi_l\widehat{f}(\xi)\\ &= 2\pi i\xi_j\Big(\sum_{k=1}^{n}\frac{\xi_k^2}{|\xi|^2}-\sum_{k<l}\frac{\xi_l\xi_k}{|\xi^2|}(e_{kl}+e_{lk})\Big)\widehat{f}(\xi) = 2\pi i\xi_j\widehat{f}(\xi)\\ &= \mathscr{F}\Big(\frac{\partial f}{\partial x_j}\Big)(\xi).\end{aligned}$$

□

4.3 The monogenic signal

Now that we have decided on a Riesz transform as an extension of the Hilbert transform, we define monogenic signals as follows.

Definition 4.3.1 (Monogenic signal)
LET $f \in L^p(\mathbb{R}^n)_n = L^p(\mathbb{R}^n, \mathbb{R}_n)$, *where* $1 < p < \infty$.
THEN *the* **monogenic signal** f_m *is defined by the following operator:*

$$\mathscr{M} : L^p(\mathbb{R}^n)_n \to L^p(\mathbb{R}^n)_n,$$
$$f \mapsto f_m = f + Rf = f + \sum_{j=1}^{n} R_j f e_j.$$

In a similar form, this has been proposed by Felsberg in his PhD-thesis [15]. However, note that the Riesz transform used to define monogenic signals in [15] is neither a left linear operator nor unitary nor self adjoint, whence the resulting monogenic signal is not left-linear. The left-linearity of the monogenic signal, however, is necessary for the proof that the monogenic signal maps frames of $L^2(\mathbb{R}^n)$ to Clifford frames of a certain Clifford-Hilbert module, a clearly favorable property. (See section 5.3.)

Remark 4.3.2

(i) LET $f \in L^2(\mathbb{R}^n)$ *be a real valued signal.*
THEN $\mathscr{M} e_0 f = f_m \in L^2(\mathbb{R}^n, \mathbb{R}^{n+1})$ *is a paravector valued function.*

(ii) The interpretation of this monogenic signal as a local Fourier transform is not possible, as there is no Fourier transform that would fit: The monogenic signal is paravector valued, while the standard Fourier transform has values in $\mathbb{C}_n$. *However, we can define an instantaneous frequency via the phase. (See section 4.4.)*

Theorem 4.3.3 (The monogenic projection)
The operator $2^{-1/2}\mathscr{M}$ *is a projection.*

Proof. Let $f \in L^p(\mathbb{R}^n)_n$. We have to show that $\mathscr{M}^2 = 2\mathscr{M}$:

$$\mathscr{M}^2 f = (\mathrm{Id}+R)(\mathrm{Id}+R)f = (\mathrm{Id}+2R+\mathrm{Id})f = 2\mathscr{M}f.$$

□

4.4 The phase of a monogenic signal

Our goal in defining a monogenic signal was to attain a decomposition of a signal into phase and amplitude. This decomposition will be the topic of the present section.

Since the monogenic signal is paravector-valued, we use Definition 3.3.7 to decompose f_m into an amplitude function a, a phase ϕ, and a vector-valued phase direction d.

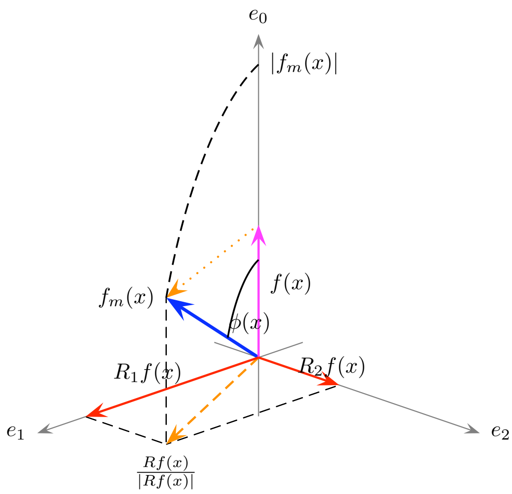


Figure 4.1: Phase $\phi(x)$ and phase direction $d(x) = \frac{Rf(x)}{|Rf(x)|}$ (direction of the dashed vector) of the 3-D vector $f_m(x) = (f(x), R_1 f(x), R_2 f(x))$.

Definition 4.4.1 (Amplitude-phase decomposition of the monogenic signal)
LET $f \in L^p(\mathbb{R}^n)$ *and let* $f_m \in L^p(\mathbb{R}^n, \mathbb{R}^{n+1})$ *be its associated monogenic signal.*

THEN *a monogenic signal can be decomposed as follows.*

$$\begin{aligned} f_m &= |f_m| \left(\frac{f}{|f_m|} + \frac{Rf}{|Rf|} \frac{|Rf|}{|f_m|} \right) \\ &= a\Big(\cos(\phi) + d \sin(\phi)\Big) = a \exp(d\phi). \end{aligned} \tag{4.1}$$

Here, $a = |f_m| \in \mathbb{R}_0^+$ *is the* **amplitude** , $\phi = \arccos\left(\frac{f}{|Rf|}\right) \in [0, \pi[$ *is the* **phase**, *and* $d = \frac{\overrightarrow{f_m}}{|\overrightarrow{f_m}|} = \frac{Rf}{|Rf|}$ *is the vector-valued* **phase direction**.

Definition 4.4.2 (Instantaneous frequency)
LET $f \in L^p(\mathbb{R}^n)$ *have the property that the phase* ϕ *is a differentiable function.*

THEN *the* **instantaneous frequency** $\omega(x)$ *is defined as the directional derivative in the direction* $d(x)$ *of the phase angle* $\phi(x)$ *at any point* $x \in \mathbb{R}^n$ *for which* $f(x) \neq 0$*:*

$$\nabla\phi(x) = \left(\frac{d\phi(x)}{dx_1}, \ldots, \frac{d\phi(x)}{dx_n} \right).$$

Theorem 4.4.3 (Computation of the instantaneous frequency)
LET f *be as in Definition 4.4.2.*
THEN

$$\omega(x) = \left\langle \nabla\phi(x), d(x) \right\rangle_{\mathbb{R}^n} = \left\langle \frac{|Rf(x)|\nabla f(x) - (\nabla|Rf(x)|)f(x)}{a^2(x)}, d(x) \right\rangle_{\mathbb{R}^n}.$$

Proof. Since $\tan(\phi(x)) = \frac{|Rf(x)|}{f(x)}$, for all $\alpha \in \{1, \ldots, n\}$, $x \in \mathbb{R}^n$, and $f(x) \neq 0$ we have that

$$\begin{aligned}
\frac{d\phi(x)}{dx_\alpha} &= \frac{d}{dx_\alpha} \arctan\left(\frac{f(x)}{|Rf(x)|}\right) \\
&= \frac{1}{1 + \left(\frac{f(x)}{|Rf(x)|}\right)^2} \frac{f(x)\frac{d}{dx_\alpha}|Rf(x)| - |Rf(x)|\frac{d}{dx_\alpha}f(x)}{f^2(x)} \\
&= \frac{f(x)\frac{d}{dx_\alpha}|Rf(x)| - |Rf(x)|\frac{d}{dx_\alpha}f(x)}{f^2(x) + \left(Rf(x)\right)^2}.
\end{aligned}$$

Hence,

$$\omega(x) = \left\langle \nabla\phi(x), d(x) \right\rangle_{\mathbb{R}^n} = \left\langle \frac{|Rf(x)|\nabla f(x) - (\nabla|Rf(x)|)f(x)}{a^2(x)}, d(x) \right\rangle_{\mathbb{R}^n}.$$

□

4.5 The monogenic signal and Cauchy transforms

Theorem 4.3.3 states that a monogenic signal is the scalar multiple of a projection. In this section we determine the space that is the range of this projection.

This yields a relation between monogenic signals and some topics in (hyper-) complex analysis. The results presented here are not new but they are important in obtaining a complete view of the monogenic signal.

We start by deriving a relation between monogenic signals, Hardy spaces, and the Cauchy, Poisson, and Riesz transforms.

Definition 4.5.1 (Hardy space)
LET $p > 0$.
THEN *the* **Hardy space**

$$H^p(\mathbb{R}^{n+1}_+) = H^p(\mathbb{R}^{n+1}_+, \mathbb{R}^{n+1})$$

is the space of functions $u : \mathbb{R}^{n+1}_+ \to \mathbb{R}^{n+1}$, *that are monogenic in the upper half plane*

$$\mathbb{R}^{n+1}_+ := \{(x_0,\ x) : x_0 \in \mathbb{R}_+, x \in \mathbb{R}^n\},$$

and satisfy

$$\|u\|_{H^p} = \sup_{x_0>0}\left(\int_{\mathbb{R}^n} |u(x_0, x)|^p dx\right)^{1/p} < \infty.$$

Theorem 4.5.2 (Monogenic signals and conjugate harmonic functions)
LET $f = \sum_{\alpha=0}^n e_\alpha f_\alpha \in L^p(\mathbb{R}^n, \mathbb{R}^{n+1})$, $1 < p < \infty$, *and let*

$$u_\alpha(x, x_0) := P_{x_0} * f_\alpha(x),\ \forall \alpha = 0, \ldots, n,$$

where P_{x_0} is the Poisson kernel defined in Definition 2.1.5 on page 23:

$$P_{x_0}(x) := c_n \frac{x_0}{(|x|^2 + x_0^2)^{\frac{n+1}{2}}},$$

where

$$c_n = \frac{\Gamma(\frac{n+1}{2})}{\pi^{\frac{n+1}{2}}},\ l = 1, \ldots, n.$$

THEN

$$f_\alpha = R_\alpha(f_0),\ \forall \alpha = 1, \ldots, n,$$

if and only if *$u \in H^p(\mathbb{R}^{n+1}_+)$. That is, u is a monogenic function in the upper half-space and thus satisfies the generalized Cauchy Riemann equations* (3.4)

$$\frac{\partial u_\alpha}{\partial x_\beta}(x, x_0) = \frac{\partial u_\beta}{\partial x_\alpha}(x, x_0),\ \forall \alpha, \beta = 1, \ldots, n,$$
$$\frac{\partial u_\alpha}{\partial x_0}(x, x_0) + \frac{\partial u_0}{\partial x_\alpha}(x, x_0) = 0,\ \forall \alpha = 1, \ldots, n,$$
$$\sum_{\alpha=1}^{n} \frac{\partial u_\alpha}{\partial x_\alpha}(x, x_0) = \frac{\partial u_0}{\partial x_0}(x, x_0),\ \forall (x, x_0) \in \mathbb{R}^{n+1}_+.$$

Proof. Suppose that $f_\alpha = R_\alpha f_0$. Then $\widehat{f}_\alpha(t) = \frac{i t_\alpha}{|t|} \widehat{f}_0(t)$ and, hence,

$$u_\alpha(x, x_0) = \int_{\mathbb{R}^n} \widehat{f}_0(t) \frac{i t_\alpha}{|t|} e^{-2\pi|t|x_0} e^{2\pi i t x} dt,\ \forall \alpha = 1, \ldots, n,$$
$$u_0(x, x_0) = \int_{\mathbb{R}^n} \widehat{f}_0(t) e^{-2\pi|t|x_0} e^{2\pi i t x} dt.$$

By the dominated convergence theorem [41], we may differentiate under the integral sign, and thus obtain

$$\frac{\partial u_0}{\partial x_0}(x, x_0) = -2\pi \int_{\mathbb{R}^n} \widehat{f}_0(t)|t| e^{-2\pi|t|x_0} e^{2\pi i t x} dt,$$
$$\frac{\partial u_0}{\partial x_\beta}(x, x_0) = 2\pi i \int_{\mathbb{R}^n} \widehat{f}_0(t) t_\beta e^{-2\pi|t|x_0} e^{2\pi i t x} dt,$$
$$\frac{\partial u_\beta}{\partial x_0}(x, x_0) = 2\pi i \int_{\mathbb{R}^n} \widehat{f}_0(t) t_\alpha e^{-2\pi|t|x_0} e^{2\pi i t x} dt,$$
$$\frac{\partial u_\alpha}{\partial x_\beta}(x, x_0) = 2\pi \int_{\mathbb{R}^n} \widehat{f}_0(t) \frac{t_\alpha t_\beta}{|t|} e^{-2\pi|t|x_0} e^{2\pi i t x} dt.$$

Now (3.4) is easy to check. The first equation follows from

$$\frac{\partial u_\alpha}{\partial x_\beta}(x, x_0) = -2\pi \int_{\mathbb{R}^n} \widehat{f}(t) \frac{t_\alpha t_\beta}{|t|} e^{-2\pi|t|x_0} e^{2\pi i t x} dt$$
$$= \frac{\partial u_\beta}{\partial x_\alpha}(x, x_0),$$

and the second identity from

$$\frac{\partial u_0}{\partial x_\beta}(x, x_0) = 2\pi i \int_{\mathbb{R}^n} \widehat{f}(t) t_\beta e^{-2\pi|t|x_0} e^{2\pi i t x} dt$$
$$= -\frac{\partial u_\beta}{\partial x_0}(x, x_0).$$

Finally, the last identity follows from

$$\begin{aligned}
\sum_{\alpha=1}^{n} \frac{\partial u_\alpha}{\partial x_\alpha}(x,x_0) &= -\sum_{\alpha=1}^{n} 2\pi \int_{\mathbb{R}^n} \widehat{f}(t)\frac{t_\alpha^2}{|t|} e^{-2\pi|t|x_0} e^{2\pi i t x} dt \\
- &= 2\pi \int_{\mathbb{R}^n} \widehat{f}(t)\frac{\sum_{\alpha=1}^{n} t_\alpha^2}{|t|} e^{-2\pi|t|x_0} e^{2\pi i t x} dt \\
&= -2\pi \int_{\mathbb{R}^n} \widehat{f}(t)|t| e^{-2\pi|t|x_0} e^{2\pi i t x} dt \\
&= \frac{\partial u}{\partial x_0}(x,x_0).
\end{aligned}$$

Conversely, let $\beta \in \{1,\dots,n\}$ and $u_\beta(x,x_0) = \int_{\mathbb{R}^n} \widehat{f_\beta}(t) e^{-2\pi|t|x_0} e^{2\pi i t x} dt$. The fact that $\frac{\partial u_0}{\partial x_\beta} = -\frac{\partial u_\beta}{\partial x_0}$, implies that

$$2\pi i \int_{\mathbb{R}^n} \widehat{f}(t)\, t_\beta e^{-2\pi|t|x_0} e^{2\pi i t x} dt = 2\pi \int_{\mathbb{R}^n} \widehat{f_\beta}(t)|t| e^{-2\pi|t|x_0} e^{2\pi i t x} dt;$$

Therefore $\widehat{f_\beta}(t) = \frac{i t_\beta}{|t|}\widehat{f}(t)$, and thus

$$f_\beta = R_\beta(f),\ \beta = 1,\dots,n.$$

□

Remark 4.5.3
Let $1 < p < \infty$. We have a one to one correspondence between functions in L^p, monogenic signals and functions on Hardy spaces in the upper half-space. As we have seen in Theorem 4.5.2, the Poisson transform of a monogenic signal, which is given by convolution with the Poisson kernel, is a monogenic function in the upper half-space $\mathbb{R}^{n+1}$.

Let $1 < p < \infty$. For every $f \in L^p$ the function $u : \mathbb{R}_+^{n+1} \to \mathbb{R}^n$, $(x_0,x) \mapsto P_{x_0} f_m(x)$ is an element of $H^p(\mathbb{R}_+^{n+1})$.

In the case $p = 1$, the additional condition $Rf \in L^1(\mathbb{R}^n,\mathbb{R}^n)$ yields $u \in H^1(\mathbb{R}_+^{n+1},\mathbb{R}^{n+1})$.

Using the **conjugate Poisson kernels**

$$q_l(x,x_0) := c_n \frac{x_l}{(|x|^2 + x_0^2)^{\frac{n+1}{2}}} = P_{x_0} R_l(x),$$

we can write u in the form

$$u(x,x_0) := u_0(x,x_0) + \sum_l e_l u_l(x,x_0) := e_0 f * P_{x_0}(x) + \sum_{l=1}^{n} e_l f * q_l(x,x_0).$$

The monogenic function equals the limit

$$\lim_{x_0 \to 0} u(x_0,x) = f_m(x),\ \text{a.e.}.$$

Furthermore, $\lim_{x_0\to 0}\int_{\mathbb{R}^n} |u(x_0,x) - f(x)|^p dx = 0$. (A proof for this can be found in [45], ChapterIII.2, Theorem 1.)

We have shown that the monogenic signal of a function in L^p is mapped by the Poisson transform to a function in the Hardy space H^p. The next theorem gives a result in the opposite direction, namely when a function in H^p has a boundary value that is a function in L^p. To state this theorem, we need the notion of non-tangential limits.

Definition 4.5.4 (Non-tangential limit)
LET $c > 0$. *The cone based at $x \in \mathbb{R}^n$ with aperture c is given by*

$$\Gamma_c(x) := \{(y_0, y) \in \mathbb{R}^{n+1}_+ : |y - x| < cx_0\}.$$

A function F on $\mathbb{R}^{n+1}_+$ has a **non-tangential limit** *at x, if*

$$\operatorname*{lim\text{-}n.t.}_{z\to x} F(z) = \lim_{\substack{z\to x \\ z\in\Gamma_c(x)}} F(z),$$

exists for some $c > 0$.

Theorem 4.5.5
LET $p > \frac{n-2}{n-1}$ *and $u \in H^p(\mathbb{R}^{n+1}_+)$.*
THEN *there exists an $f \in L^p(\mathbb{R}^n)$ such that*

1. $\text{lim-n.t.}_{z\to x}\, u(z) = f(x)$, *a.e.;*
2. $\lim_{x_0\to 0} \int_{\mathbb{R}^n} |u(x_0, x) - f(x)|^p dx = 0$.

Proof. See [19], chapter 2, section 5. (Note the different convention with respect to the sign of the Riesz transform used there.) □

What remains to be shown is that the limit of the function u is again a monogenic signal. This is proved in Theorem 4.5.2.

However, this result also be derived using the Cauchy integral of Clifford analysis.

Definition 4.5.6 (Cauchy Integral)
LET $1 \le p < \infty$, *and $f \in L^p(\mathbb{R}^n, \mathbb{R}^{n+1})$. Let $z \in \mathbb{R}^{n+1} \setminus \mathbb{R}^n$, i.e., z is an element of the upper or lower half-space; specifically $\langle z\rangle_0 \neq 0$.*
THEN *the* **Cauchy integral** *is defined by*

$$Cf(z) := \frac{1}{c_n} \int_{\mathbb{R}^n} \frac{u - z}{|u - z|^n} n(u) f(u) du,$$

where $n(x) = -e_0$ is the outward pointing normal vector to $\mathbb{R}^{n+1}_+$ which is constant.

Note that Cf is monogenic on $\mathbb{R}^{n+1}_+$. Furthermore the Cauchy integral is closely related to monogenic signals: $Cf(x_0, x) = \frac{1}{2}\big(P_{x_0} * f(x) + \sum_{l=1}^n e_l q_l(x_0, \cdot) * f(x)\big) = \frac{1}{2} P_{x_0} f_m(x)$, $\forall x_0 > 0$.

Theorem 4.5.7 (Cauchy integral and monogenic signals)
LET *either* $1 < p < \infty$ *and* $f \in L^p(\mathbb{R}^n, \mathbb{R}^{n+1})$, *or* $p = 1$ *and* $f \in L^1(\mathbb{R}^n, \mathbb{R}^{n+1}), Rf \in L^1(\mathbb{R}^n, \mathbb{R}^{n+1})$.
THEN $Cf \in H^p(\mathbb{R}^{n+1}_+, \mathbb{R}^{n+1})$ *and*

$$\operatorname*{lim\text{-}n.t.}_{z\to x} Cf(z) = \frac{1}{2} f_m(x), \text{ a.e..}$$

Proof. See [19], chapter 2, section 5. □

Theorem 4.5.8
LET $1 \le p < \infty$ *and* $u \in H^p(\mathbb{R}_+^{n+1}, \mathbb{R}^{n+1})$. *Denote* $\mu := \text{lim-n.t.}_{z\to\cdot}\, u(z)$.
THEN

$$u = C\mu.$$

Proof. See [19], chapter 2, section 5. □

Remark 4.5.9 (The Cauchy integral in the lower half-space)
There is an analogous theory for Hardy spaces of the lower half-space. Let $f \in L^p(\mathbb{R}^n, \mathbb{R}^{n+1})$, *where* $1 \le p < \infty$. *Let* $z \in \mathbb{R}^{n+1} \setminus \mathbb{R}^n$, *i.e.,* z *is an element of the lower half-space; specifically* $\langle z \rangle_0 \ne 0$ *and let* $n(u) = e_0$ *be the outward pointing normal vector of* $\mathbb{R}^n \subset \mathbb{R}^{n+1}$. *If we let* $H^p(\mathbb{R}_-^{n+1}, \mathbb{R}^{n+1})$ *be the Hardy space of the lower half-space, then the Cauchy integral*

$$\overline{C}f(z) := \frac{1}{c_n}\int_{\mathbb{R}^n} \frac{z-u}{|u-z|^n}\, n(u) f(u)\, du = \frac{1}{2}\Big(P_{x_0} * f(x) - \sum_{l=1}^{n} e_l q_l(x_0, \,\cdot\,) * f(x)\Big)$$

has the following properties.
Let either $1 < p < \infty$ *and* $f \in L^p(\mathbb{R}^n, \mathbb{R}^{n+1})$, *or* $p = 1$ *and* $f, Rf \in L^p(\mathbb{R}^n, \mathbb{R}^{n+1})$.
Then $\overline{C}f \in H^p(\mathbb{R}_-^{n+1}, \mathbb{R}^{n+1})$, *and*

$$\underset{z\to x}{\text{lim-n.t.}}\, \overline{C}f(z) = \frac{1}{2}f(x) - Rf(x), \text{ a.e..}$$

Let $1 \le p < \infty$ *and* $u \in H^p(\mathbb{R}_-^{n+1}, \mathbb{R}^{n+1})$.
Then

$$u = \overline{C}(\underset{z\to\cdot}{\text{lim-n.t.}}\, u(z)).$$

Remark 4.5.10 (Range of a monogenic signal)
The non tangential limits of Cauchy integrals define singular operators. These operators

$$P_\pm : L^2(\mathbb{R}^n) \to L^2(\mathbb{R}^n),\ f \mapsto \frac{1}{2}(f \pm Rf)$$

are projections called **Plemelj projections**. *They satisfy*

$$P_+P_- = P_-P_+ = 0, \quad P_\pm^2 = P_\pm, \text{ and } P_+ + P_- = \text{Id}.$$

Furthermore, the Cauchy integral is the Poisson transform of the corresponding Plemelj projection.
Let $1 < p < \infty$, *then*

$$TH_\pm^p(\mathbb{R}^n) := \{f \pm Rf : f \in L^p(\mathbb{R}^n, \mathbb{R}^{n+1})\},$$

and, for $p = 1$,

$$TH_\pm^1(\mathbb{R}^n) := \{f \pm Rf : f, Rf \in L^1(\mathbb{R}^n, \mathbb{R}^{n+1})\}.$$

$P_\pm$ *are obviously projections onto* $TH_\pm^p(\mathbb{R}^n)$. *Furthermore,*

$$\mathcal{M}f = f_m = 2P_+f,\ \forall f \in TH_+(\mathbb{R}^n).$$

Let $1 \le p < \infty$. *Then the Poisson transform is an isomorphism between* $H^p(\mathbb{R}_\pm^{n+1}, \mathbb{R}^{n+1})$ *and* $TH_\pm^p(\mathbb{R}^n)$. *Moreover, for* $1 < p < \infty$, *the operator*

$$\mathcal{M}_\pm : L^p(\mathbb{R}^n, \mathbb{R}) \to TH_\pm^p(\mathbb{R}^n),\ f \mapsto (f \pm Rf)$$

is an isomorphism with inverse

$$\mathcal{M}_{\pm}^{-1} : TH_{\pm}^{p}(\mathbb{R}^n) \to L^p(\mathbb{R}^n,\mathbb{R}),\ f \mapsto \Re f.$$

Let $1 \le p < \infty$. Then the Cauchy integral is an isomorphism between $L^p(\mathbb{R}^n)$ and $H^p(\mathbb{R}_{\pm}^{n+1},\mathbb{R}^{n+1})$. In addition, the Cauchy integral is an isomorphism between $TH^p(\mathbb{R}^n)$ and $H^p(\mathbb{R}_{\pm}^{n+1},\mathbb{R}^{n+1})$.

To summarize, we give a list of the operators used in this section:

monogenic signal	$\mathcal{M}$:	$L^p(\mathbb{R}^n) \to TH_{+}^{p}(\mathbb{R}^n)$,	$f \mapsto f_m$;
real part	$\Re$:	$TH_{+}^{p}(\mathbb{R}^n) \to L^p(\mathbb{R}^n)$,	$f_m \mapsto f$;
Poisson transform	P :	$TH_{+}^{p}(\mathbb{R}^n) \to H^p(\mathbb{R}_{+}^{n+1},\mathbb{R}^{n+1})$,	$f_m \mapsto C_f$;
Cauchy transform	C :	$L^p(\mathbb{R}^n) \to H^p(\mathbb{R}_{+}^{n+1},\mathbb{R}^{n+1})$,	$f \mapsto Cf$;
tangential limes	lim-n.t. :	$H^p(\mathbb{R}_{+}^{n+1},\mathbb{R}^{n+1}) \to TH_{+}^{p}(\mathbb{R}^n)$,	$Cf \mapsto f_m$.

4.6 Alternative analytical signals

An analytical signal is uniquely defined in two different ways: On the one hand, the Poisson transform of an analytical signal satisfies the Cauchy Riemann equations as stated in Theorem 4.5.2. On the other hand, an analytical signal is defined by the Hilbert transform, which is is uniquely defined modulo a multiplicative constant by the following four properties stated in Theorem 1.4.2. The Hilbert transform

(i) anti-commutes with reflection, i.e. $\mathcal{H}(f(-\,\cdot\,))(x) = -\mathcal{H}(f)(-x)$;

(ii) commutes with translation;

(iii) commutes with dilation;

(iv) $\exists A(p)$ such that $\|\mathcal{H}f\|_p \le A(p)\|f\|_p \quad \forall f \in L^p(\mathbb{R})$.

Properties (ii) and (iv) are extended to higher dimensions in the obvious way. There are however different approaches to extend properties (i) and (iii) to higher dimensions. Each of these approaches yields a uniquely defined extension of the Hilbert transform. The important property here is property (i), which states the behaviour of the Hilbert transform under a group of symmetries, i.e. the reflections. Dilation is then determined by the condition that it commutes with these symmetries.

Remark 4.6.1 (Uniqueness of the Riesz transform)
For the definition of the Riesz transform we chose as symmetries the group of reflections and rotations in $\mathbb{R}^n$ given by the matrix group $O(n)$. As a consequence we extend property (iii) by isotropic dilations and property (i) by the corresponding property with respect to rotations and reflections in $\mathbb{R}^n$ stated in Theorem 2.1.2(i). This yields the steerability of the Riesz transform Definition 2.1.9. Together with properties (ii) and (iv) the Riesz transform is uniquely defined.

The identity of objects in images is not changed by isotropic dilations and rotations. Hence for image analysis it is important that the action of rotation on images can be controlled. This is just what steerability of the Riesz transform means. Property (i) states that Riesz transforms are a representation of the rotation group. That is, the Riesz transform originates from symmetries which are exactly those needed for image processing.

There are other representations of the rotation group on which the higher Riesz transforms in chapter 8 are based.

We have seen in Theorem 4.5.2 that the Poisson transform of hypercomplex monogenic signals satisfies the Cauchy-Riemann equations as stated in Theorem 4.5.2. This is another aspect under which the monogenic signal is a unique extension of the analytical signal.

The Riesz transform is not invariant under dilations of the form

$$\mathscr{D}_{a_1,\dots,a_n} = \mathrm{diag}(a_1,\dots,a_n). \tag{4.2}$$

Thus a different transform is needed to define an analytical signal in this setting. The partial Hilbert transform with axis x_α, $\alpha = 1,\dots,n$, is given by the Fourier multiplier

$$\widehat{\mathscr{H}_\alpha} := \frac{i x_\alpha}{|x_\alpha|}.$$

Partial Hilbert transforms are invariant under dilations of the form (4.2). Hence the transform of choice is the combination of all partial Hilbert transforms:

$$\mathscr{H} f = \sum_{\boldsymbol{\alpha}\in\mathscr{O}_n} e_{\boldsymbol{\alpha}} \mathscr{H}_{\boldsymbol{\alpha}} f,\ e_{\boldsymbol{\alpha}} \in \mathbb{R}_n,$$

where $\boldsymbol{\alpha} = (\alpha_1,\dots,\alpha_{|\boldsymbol{\alpha}|})$ is a multiindex and $\mathscr{H}_{\boldsymbol{\alpha}} = \mathscr{H}_{\boldsymbol{\alpha}_1}\dots\mathscr{H}_{\boldsymbol{\alpha}_{|\boldsymbol{\alpha}|}}$. However, these partial Hilbert transforms do not commute with rotations. The correspondence with property (i) is the fact that partial Hilbert transforms anti-commute with reflections in the plane perpendicular to the axis of the partial Hilbert transform, and commute with all reflections in planes that contain the axis of the partial Hilbert transform. In the two dimensional case this is the Hilbert transform proposed by Bülow in his thesis [8] and Baraniuck in [9].

This definition of an multidimensional analytical signal can be interpreted as an analytical signal where the term analytical corresponds to a different kind of Cauchy-Riemann like equations.

In $\mathbb{R}^n \times \mathbb{R}_+$, take the Poisson kernel $p(x,y) = \frac{y}{\pi\|x+y\|^2}$ and its conjugate $q(x,y) = \frac{x}{\pi\|x+y\|^2}$.
The analytical function of a 2-D function corresponding to the Quaternion Wavelet Transform of Baraniuck [9] found in the PhD-thesis of Bülow´s [8] would be $u + i v_i + j v_j + k v_k$, where

$$\begin{aligned} u &:= f * p(x_1,y_1) * p(x_2,y_2);\\ v_i &:= f * q(x_1,y_1) * p(x_2,y_2);\\ v_j &:= f * p(x_1,y_1) * q(x_2,y_2);\\ v_k &:= f * q(x_1,y_1) * q(x_2,y_2). \end{aligned}$$

Thus, the analytical extension can be written as

$$(u + i v_i) + j(v_j + i v_k) := u_2 + j v_2 = u_1 + i v_1 =: (u + j v_j) + i(v_i + j v_k),$$

satisfying the Cauchy-Riemann-like equations given in [34]:

$$\begin{aligned} \frac{\partial u_l}{\partial y_l} + \frac{\partial v_l}{\partial x_l} &= 0,\\ \frac{\partial v_l}{\partial y_l} + \frac{\partial u_l}{\partial x_l} &= 0,\ l = i,j. \end{aligned}$$

Chapter 5

Monogenic wavelets

The agenda of this chapter is to implement the monogenic signal defined in the last chapter via wavelet frames.

The first section 5.1 motivates the use of wavelet frames for the implementation. Example 5.1.2 illustrates our approach using the example of analytical wavelet frames and sets the roadmap for sections 2 and 3.

Section 5.2 extends the concept of frames on Hilbert spaces to that of Clifford-Hilbert modules – the space in which the monogenic signal lives – introduced in section 3.2. Our concept of Clifford frames – which use Clifford algebra valued frame coefficients to achieve a frame decomposition – appears to be entirely novel and it is especially remarkable that Theorem 5.2.3 holds although the Cauchy-Schwartz inequality is not valid for the Clifford algebra valued inner product $\langle\langle \cdot, \cdot \rangle\rangle$.

Section 5.3 defines the notion of a hypercomplex wavelet transform, which is compatible with Clifford frames (Definition 5.3.1) and shows that we can derive a monogenic wavelet frame from a wavelet frame for $L^2(\mathbb{R}^n)$ (Theorem 5.3.3). Furthermore, Theorem 5.4.3 states that the decay rate of the wavelet frames is preserved by the Riesz transform.

In section 5.4 we derive novel conditions under which the partial Riesz transform of a function inherits the decay rate.

Section 5.5 is dedicated to the search for suitable wavelet frames. First, we introduce results which reduce the problem of constructing frames to the problem of constructing Riesz partitions of unity. Theorem 5.5.7 provides a new way to find such Riesz partitions of unity. An explicit construction is then given in Example 5.5.1. Theorem 5.5.8 shows how to derive Riesz partitions of unity from given compactly supported wavelet orthonormal bases of $L^2(\mathbb{R})$.

Finally section 5.6 gives the implementation of the wavelet frames of Example 5.5.1 as an imageJ plugin. The software and the applications have been the topic of the diploma thesis of Martin Storath [47] under supervision of the author.

Some of the results of this chapter have been published in the peer reviewed article [24] and in [23].

5.1 Introduction and example

In the last chapter we showed that monogenic signals based on the Riesz transform are the natural extension of analytical signals to arbitrary finite dimension in the context of image processing.

Hence it is time to think about the implementation of the monogenic signal. The way we want to do this is by using monogenic wavelets. There are three main reasons to use wavelets here:

The first reason to use wavelets is that the multi-scale phase and amplitude of the wavelets are more useful than the phase of the monogenic signal, which can be interpreted as the average of the phase over all scales. To demonstrate this let us consider an example.

EXAMPLE 5.1.1 (Phase separation via wavelets):
LET

$$f(t) = a(t)\cos(\omega t),$$

where $a \in L^2(\mathbb{R}) : \operatorname{supp}(\widehat{a}) \in]-\omega, \omega[$.

THEN by the Bedrosian identity Theorem 2.2.16 on page 36 its analytical signal is $a(t)e^{i\omega t}$. The amplitude of this function is $a(t)$, its phase is ωt, and the instantaneous frequency is ω.

LET us now consider the function

$$g(t) = a(t)(\cos(\omega t) + \cos(\gamma t)),$$

where $0 < \omega < \gamma < \infty$.

THEN the amplitude of g is $a(t)|\cos(\frac{\omega-\gamma}{2}t)|$, its phase is $\frac{\omega+\gamma}{2}t$ and the instantaneous frequency is $\frac{\omega+\gamma}{2}$. That is the phase is an average of the two phases which are present. The amplitude clearly has some component which we would consider to be part of the phase.

LET $\{2^j T_k \psi\}_{j,k\in\mathbb{Z}}$ be a wavelet set such that $\operatorname{supp}(\widehat{\psi}) \subseteq [-1/4, -1/8] \cup [1/8, 1/4]$ and let $\phi \geq \omega$. By Theorem 1.3.15 on page 10 we know that the Fourier transform of $a(t)\cos(\phi t)$ is supported on $\phi + \operatorname{supp}(\widehat{a})$. Let $d := \sup\left\{|x-y| : x, y \in \operatorname{supp}(\widehat{a})\right\}$, $a_\phi := \lfloor \log_2(\phi - d) \rfloor$ and $b_\phi := \lceil \log_2(\phi + d) \rceil$.

THEN it follows that $\operatorname{supp}\left(\mathscr{F}\left(a\cos(\phi\,\cdot\,)\right)\right) \subseteq [2^{a_\phi}, 2^{b_\phi}]$ and if $b_\omega < a_\gamma$ then the wavelet transform separates the two components which may then be analysed separately as shown in Example 5.1.2. Since the analytical wavelet coefficients of $a\cos(\phi)$ have disjoint support with respect to scale from those of $a\cos(\omega)$ the phase of these two signals is well separated by the analytical wavelet.

The second reason to use wavelets for the implementation of the Riesz transform is that the Fourier multiplier of the Riesz transform is discontinuous at 0. Thus the Riesz transform of a signal in $L^1(\mathbb{R}^n)$ is not an element of $L^1(\mathbb{R}^n)$ if its 0-th moment does not vanish. In fact, fast decay of the Riesz transform of a signal requires vanishing moments as shown in Theorem 5.4.3. As a consequence we will use wavelets with vanishing moments to implement the Riesz transform.

The third reason is the fact that the Riesz transform behaves remarkably well under the operators which generate the wavelet system from the mother wavelet. Indeed, the Riesz transform is uniquely defined by the properties that it commutes with translation and isotropic dilation and behaves like a vector under rotation (Theorem 2.1.2 on page 20). That is Theorem 2.1.2(i) yields an irreducible unitary representation of the rotation group. Hence the dual frame of the Riesz transformed wavelet frame is the Riesz transform of the dual wavelet frame as shown in Theorem 5.3.4.

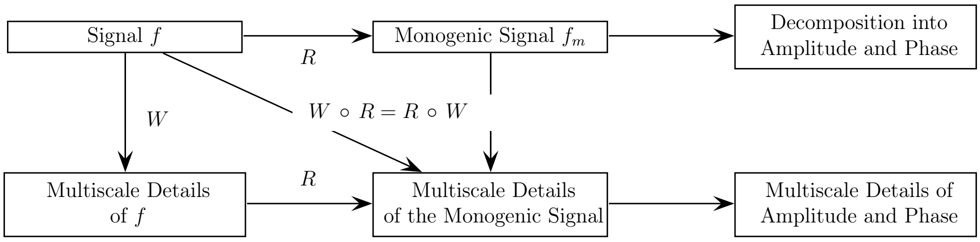


Figure 5.1: Commutative diagram for the Riesz (R) and wavelet (W) transform. For a multiscale decomposition of the details of amplitude and phase we can use either $R \circ W$ or $W \circ R$, since they are equal.

The example of analytical wavelets in the $1-D$ case will demonstrate how the implementation via wavelets works.

EXAMPLE 5.1.2 (Analytical wavelets):
LET $\psi \in L^2(\mathbb{R})$ be a mother wavelet for a wavelet frame.

THEN $\mathcal{H}\psi$, the Hilbert transform of ψ, is a mother wavelet that generates a wavelet frame for $L^2(\mathbb{R},\mathbb{C})$. Furthermore, ψ_a the analytical signal of ψ, is a mother wavelet for a wavelet frame of $TH^2_-(\mathbb{R})$ – the space of tangential limits of functions in the Hardy space – with complex frame coefficients.

LET $f \in L^2(\mathbb{R})$ and let f_a be its analytical signal.

THEN

$$W_{\psi_a} f = W_\psi f_a$$

and hence we can compute the analytical signal of f via

$$f_a = W_\psi^{-1} W_{\psi_a} f.$$

Proof. The Hilbert transform is a bounded invertible mapping in $L^2(\mathbb{R})$. It follows by Theorem 1.3.27 on page 13, that $\mathcal{H}\psi$ generates a frame for $L^2(\mathbb{R})$.

We know that the operator mapping a signal to an analytical signal to be bounded and surjective onto $TH^2_-(\mathbb{R})$. Hence Theorem 1.3.27 shows that ψ_a generates a frame for $TH^2_-(\mathbb{R})$.

The wavelet transform of f is

$$\begin{aligned}
W_{\psi_a} f(d,t) &= \langle D_d T_t \psi_a, f\rangle \\
&= \langle D_d T_t(\psi + i\mathcal{H}\psi), f\rangle \\
&= \langle D_d T_t \psi, f\rangle + i\langle \mathcal{H} D_d T_t \psi, f\rangle \\
&= \langle D_d T_t \psi, f\rangle + \int_{\mathbb{R}} \operatorname{sgn}(\xi) \widehat{D}_d T_t \psi(\xi) \overline{\widehat{f}(\xi)} d\xi \\
&= \langle D_d T_t \psi, f\rangle - i\langle D_d T_t \psi, \mathcal{H} f\rangle \\
&= W_\psi f_a(d,t).
\end{aligned}$$

Here we used the fact that the Hilbert transform commutes with translation and dilation Theorem 1.4.2 on page 15 and Parsevals formula Corollary 1.3.14 on page 10. □

5.2 Frames on Clifford-Hilbert modules

We wish to extend the analytical wavelets of Example 5.1.2 to our monogenic signal Definition 4.3.1. Hence we need to define wavelet frames on Clifford-Hilbert modules. That is we need a concept of frames in Clifford-Hilbert modules with frame coefficients in the Clifford algebra. These frames will be called **Clifford frames**. They will be generated using an analysis operator which is based on the Clifford algebra valued product $\langle\!\langle \cdot , \cdot \rangle\!\rangle$ rather than on the scalar product $(\cdot , \cdot)$. (See subsection 3.2.2 for the definitions of $\langle\!\langle \cdot , \cdot \rangle\!\rangle$ and $(\cdot , \cdot)$.)

In the following we follow the general procedure for Hilbert spaces in [10], Chapters 3.2 and 5, modify and construct new proofs for the case of Clifford-Hilbert modules when necessary. It is possible to roughly follow the proofs of the Hilbert space case because the Hilbert module $\left(H_n, \langle\!\langle \cdot,\cdot \rangle\!\rangle\right)$ over the Hilbert space H is in itself a Hilbert space $\left(H_n, (\cdot,\cdot)\right)$.

Difficulties in adapting the proofs are due to three facts:

- The coefficients of the Clifford frame are not scalars with respect to the Hilbert space $\left(H_n, (\cdot,\cdot)\right)$.
- $\langle\!\langle \cdot , \cdot \rangle\!\rangle$ is not a scalar product.
- The Clifford algebra setting we consider is not commutative.

Considering the modifications in the Cauchy Schwartz inequality Theorem 3.2.4(iii) on page 50 it is especially remarkable that Theorem 5.2.3 holds with the same bounds as for Bessel sequences in Hilbert spaces.

An outstanding feature of the Clifford frames is that they give a unified framework for frames and multi-frames. (See Theorem 5.3.3 for an example.)

In the following $\left(H, \langle \cdot , \cdot \rangle\right)$ will be a Hilbert space and $\left(H_n, \langle\!\langle \cdot , \cdot \rangle\!\rangle\right)$ will be the Clifford-Hilbert module defined in subsection 3.2.2.

5.2.1 Bessel sequences

Lemma 5.2.1 (The Synthesis and the Analysis Operator)
LET $\{f_k\}_{k\in\mathbb{N}} \subset H_n$ *be a sequence in the Clifford-Hilbert module* H_n*, and suppose that* $\sum_{k\in\mathbb{N}} c_k f_k$ *is convergent for all* $\{c_k\}_{k\in\mathbb{N}} \in l_n^2(\mathbb{N}) \cong l^2(\mathbb{N}, \mathbb{C}_n)$.

THEN *the* **synthesis operator**

$$T : l_n^2(\mathbb{N}) \mapsto H_n, \quad \{c_k\} \mapsto \sum_{k\in\mathbb{N}} c_k f_k$$

defines a bounded $\mathbb{C}_n$*-left-linear operator. The adjoint, called the* **analysis operator***, is given by*

$$T^* : H_n \mapsto l_n^2(\mathbb{N}), \quad f \mapsto \left\{\langle\!\langle f, f_k \rangle\!\rangle\right\}_{k\in\mathbb{N}}.$$

Furthermore

$$\sum_{k\in\mathbb{N}} \left|\langle\!\langle f, f_k \rangle\!\rangle\right|^2 \le \|T\|^2 \|f\|^2, \ \forall f \in H_n.$$

Proof. Let $l \in \mathbb{N}$. Consider the sequence of bounded left-linear operators

$$T_l : l_n^2(\mathbb{N}) \mapsto H_n, \quad \{c_k\}_{k\in\mathbb{N}} \mapsto \sum_{k=1}^{l} c_k f_k.$$

Clearly $T_l \to T$ pointwise as $l \to \infty$, thus T is bounded. (See [56] IV.2.5, a corollary of the Banach-Steinhaus theorem.)

Now let $f \in H_n$, $\{c_k\}_{k\in\mathbb{N}} \in l_n^2(\mathbb{N})$. Then

$$\begin{aligned}\langle\langle f, T(\{c_k\}_{k\in\mathbb{N}})\rangle\rangle_{H_n} &= \langle\langle f, \sum_{k\in\mathbb{N}} c_k f_k\rangle\rangle = \sum_{k\in\mathbb{N}} \langle\langle f, f_k\rangle\rangle\overline{c_k} \\ &= \sum_{k\in\mathbb{N}} \sum_{\boldsymbol{\alpha},\boldsymbol{\beta}\in\mathcal{O}_n} e_{\boldsymbol{\alpha}}\overline{e_{\boldsymbol{\beta}}}\langle\langle\langle f, f_k\rangle\rangle\rangle_{\boldsymbol{\alpha}}\langle\overline{c_k}\rangle_{\boldsymbol{\beta}} = \langle\langle\{\langle\langle f, f_k\rangle\rangle\}_{k\in\mathbb{N}}, \{c_k\}_{k\in\mathbb{N}}\rangle\rangle_{l_n^2(\mathbb{N})}. \end{aligned} \tag{5.1}$$

Since T is bounded, T^* is a bounded linear operator $T^* : H_n \to l_n^2(\mathbb{N})$. Hence the k-th coordinate functional is bounded from $H_n \mapsto \mathbb{C}_n$. By the Riesz representation theorem (Theorem 3.2.5 on page 52) there exists a sequence $\{g_k\} \subset H_n$ such that

$$T^* f = \{\langle\langle f, g_k\rangle\rangle_{H_n}\}_{k\in\mathbb{N}}.$$

By (5.1) it follows that

$$\sum_{k\in\mathbb{N}} \langle\langle f, f_k\rangle\rangle_{H_n}\overline{c_k} = \langle\langle f, T(\{c_k\}_{k\in\mathbb{N}})\rangle\rangle_{H_n} = \langle\langle T^* f, \{c_k\}_{k\in\mathbb{N}}\rangle\rangle_{l_n^2(\mathbb{N})} = \sum_{k\in\mathbb{N}} \langle\langle f, g_k\rangle\rangle_{H_n}\overline{c_k},$$

and hence $g_k = f_k$.

Now $\sum_{k\in\mathbb{N}} |\langle\langle f, f_k\rangle\rangle|^2 = \|T^* f\|^2 \le \|T^*\|^2\|f\|^2 = \|T\|^2\|f\|^2$. □

Definition 5.2.2 (Bessel Sequence)

LET $\{f_k\}_{k\in\mathbb{N}} \subset H_n$ *be a sequence such that there exists a constant* $B > 0$ *satisfying*

$$\sum_{k\in\mathbb{N}} |\langle\langle f, f_k\rangle\rangle|^2 \le B\|f\|^2 \quad \forall f \in H_n.$$

THEN $\{f_k\}_{k\in\mathbb{N}}$ *is called a* **Bessel sequence**. *A number* B *satisfying this inequality is called a* **Bessel bound**.

Theorem 5.2.3 (Boundedness of the synthesis operatr)

LET $\{f_k\}_{k\in\mathbb{N}} \subset H_n$ *be such that the synthesis operator*

$$T : l_n^2(\mathbb{N}) \to H_n, \ \{c_k\}_{k\in\mathbb{N}} \mapsto \sum_{k\in\mathbb{N}} c_k f_k$$

is a well defined left-linear bounded operator and $\|T\| \le \sqrt{B}$.

THEN AND ONLY THEN *is* $\{f_k\}_{k\in\mathbb{N}}$ *a Bessel sequence with Bessel bound* B.

Proof. Let $\{f_k\}_{k\in\mathbb{N}} \subset H_n$ be a Bessel sequence with Bessel bound B.

We will first show that $\sum_{k\in\mathbb{N}} c_k f_k$ is convergent. Let $0 < l < m < \infty$.

$$\begin{aligned}
&\Big\| \sum_{k=1}^{m} c_k f_k - \sum_{k=1}^{l} c_k f_k \Big\| = \Big\| \sum_{k=l+1}^{m} c_k f_k \Big\| \overset{3.2.4(vii)}{\leq} \sup_{\substack{\|g\|\leq 1\\ g\in H_n}} \Big| \Big(\sum_{k=l+1}^{m} c_k f_k, g \Big) \Big| \\
&\leq \sup_{\|g\|\leq 1} \sum_{k=l+1}^{m} |(c_k f_k, g)| = \sup_{\|g\|\leq 1} \sum_{k=l+1}^{m} \big| \Re\big(c_k \langle\!\langle f_k, g \rangle\!\rangle \big) \big| \\
&= \sup_{\|g\|\leq 1} \sum_{k=l+1}^{m} \Big| \sum_{\boldsymbol{\alpha}\in\mathcal{O}_n} |e_{\boldsymbol{\alpha}}|^2 \langle c_k \rangle_{\boldsymbol{\alpha}} \Big\langle \langle\!\langle f_k, g \rangle\!\rangle \Big\rangle_{\boldsymbol{\alpha}} \Big| \\
&\overset{*}{\leq} \sup_{\|g\|\leq 1} \sum_{k=l+1}^{m} \Big(\sum_{\boldsymbol{\alpha}\in\mathcal{O}_n} |\langle c_k \rangle_{\boldsymbol{\alpha}}|^2 \Big)^{1/2} \Big(\sum_{\boldsymbol{\alpha}\in\mathcal{O}_n} |\langle \langle\!\langle f_k, g \rangle\!\rangle \rangle_{\boldsymbol{\alpha}}|^2 \Big)^{1/2} \\
&= \sup_{\|g\|\leq 1} \sum_{k=l+1}^{m} |c_k| \big| \langle\!\langle f_k, g \rangle\!\rangle \big| \leq \Big(\sum_{k=l+1}^{m} |c_k|^2 \Big)^{1/2} \sup_{\|g\|\leq 1} \Big(\sum_{k=l+1}^{m} |\langle\!\langle f_k, g \rangle\!\rangle|^2 \Big)^{1/2} \\
&\leq \sqrt{B} \Big(\sum_{k=l+1}^{m} |c_k|^2 \Big)^{1/2}
\end{aligned}$$

For (*) we used the Cauchy-Schwartz inequality in the complex Hilbert space $\big(\mathbb{C}_n, (\cdot,\cdot)\big)$.

Since $\{c_k\}_{k\in\mathbb{Z}} \in l_n^2(\mathbb{N})$ this shows that $\big\{\sum_{k=1}^{m} c_k f_k\big\}_{m=1}^{\infty}$ is a Cauchy sequence in H_n and hence convergent as a sequence in the complex Hilbert space $\big(H_n, (\cdot,\cdot)\big)$. Thus $T(\{c_k\}_{k\in\mathbb{N}})$ is well defined. T is obviously left-linear. Choosing $m = 0$ and letting $l \to \infty$ in the calculation above gives the boundedness of T.

The second statement has already been shown in Lemma 5.2.1. □

Corollary 5.2.4
LET $\{f_k\}_{k\in\mathbb{N}} \subset H_n$ *be a sequence and let* $\sum_{k\in\mathbb{N}} c_k f_k$ *be convergent for all* $\{c_k\}_{k\in\mathbb{N}} \in l_n^2(\mathbb{N})$.
THEN $\{f_k\}_{k\in\mathbb{N}}$ *is a Bessel sequence.*

Corollary 5.2.5 (Unconditional convergence of Bessel sequences)
LET $\{f_k\}_{k\in\mathbb{N}} \subset H_n$ *be a Bessel sequence.*
THEN $\sum_{k\in\mathbb{N}} c_k f_k$ *converges unconditionally for all* $\{c_k\}_{k\in\mathbb{N}} \in l_n^2(\mathbb{N})$.

Lemma 5.2.6 (Bessel sequence on a dense set)
LET $\{f_k\}_{k\in\mathbb{N}} \subset H_n$ *and suppose there exists a dense subset* $V \subset H_n$ *and a constant* $B > 0$ *such that*

$$\sum_{k\in\mathbb{N}} |\langle\!\langle f, f_k \rangle\!\rangle|^2 \leq B\|f\|^2, \ \forall f \in V.$$

THEN $\{f_k\}_{k\in\mathbb{N}}$ *is a Bessel sequence in* H_n *with bound* B.

Proof. The topology of the Clifford-Hilbert module $H_n$is defined by the Euclidean norm and thus identical to the topology on the complex Hilbert space $\big(H_n, (\cdot,\cdot)\big)$. Thus, assume

$$\exists g \in H_n : \sum_{k\in\mathbb{N}} \big|\langle\!\langle g, f_k \rangle\!\rangle\big|^2 > B\|g\|^2.$$

Then there exists a finite set

$$F \subset \mathbb{N} : \sum_{k\in F} \big|\langle\!\langle g, f_k \rangle\!\rangle\big|^2 > B\|g\|^2.$$

Since

$$|\langle\langle \cdot, f_k \rangle\rangle|^2 = \left\langle \langle\langle \cdot, f_k \rangle\rangle \langle\langle f_k, \cdot \rangle\rangle \right\rangle_0,$$

we know that the operator $g \mapsto \sum_{k\in F} |\langle\langle g, f_k \rangle\rangle|^2$ is continuous. Since V is dense in H_n this implies

$$\exists h \in V : \sum_{k\in F} |\langle\langle h, f_k \rangle\rangle|^2 > B\|h\|^2.$$

But this is a contradiction.

□

5.2.2 Clifford frames

Definition 5.2.7 (Clifford Frame)
LET $\{f_k\}_{k\in\mathbb{N}} \subset H_n$ *be a sequence such that there exist constants* $0 < A \leq B < \infty$, *satisfying the* **frame inequality**

$$A\|f\|^2 \leq \sum_{k\in\mathbb{N}} |\langle\langle f, f_k \rangle\rangle|^2 \leq B\|f\|^2, \ \forall f \in H_n. \tag{5.2}$$

THEN $\{f_k\}_{k\in\mathbb{N}}$ *is called a* **Clifford frame** *for* H_n. *A is called a* **lower frame bound**, *B is called an* **upper frame bound**.

- *A Clifford frame is called* **tight**, *iff choosing* $A = B$ *is possible.*
- *Let* T *be the synthesis operator of the frame* $\{f_k\}$. *Then the* **Clifford frame operator** *is defined by*

$$S : H_n \mapsto H_n, \quad f \mapsto Sf := TT^* f = \sum_{k\in\mathbb{N}} \langle\langle f, f_k \rangle\rangle f_k.$$

- *A sequence* $\{g_k\}_{\in\mathbb{N}} \subset H_n$ *is called a* **dual frame** *of the Clifford frame* $\{f_k\}_{k\in\mathbb{N}}$ *iff* $\{g_k\}_{k\in\mathbb{N}}$ *is a Clifford frame and the* **frame decompostion**

$$f = \sum_{k\in\mathbb{N}} \langle\langle f, f_k \rangle\rangle g_k = \sum_{k\in\mathbb{N}} \langle\langle f, g_k \rangle\rangle f_k$$

  holds for all $f \in H_n$.

Remark 5.2.8 (Frame operator and frame inequality)
For frames on Hilbert spaces the equality

$$\sum_k |\langle f, f_k \rangle|^2 = \langle Sf, f \rangle, \ \forall f \in H \tag{5.3}$$

is very useful. Note that for Clifford frames in general

$$\langle\langle Sf, f \rangle\rangle = \sum_k \langle\langle f, f_k \rangle\rangle \langle\langle f_k, f \rangle\rangle \neq \sum_k |\langle\langle f, f_k \rangle\rangle|^2.$$

For Clifford frames equation (5.3) *is replaced by*

$$\sum_k |\langle\langle f, f_k \rangle\rangle|^2 = (Sf, f), \ \forall f \in H_n.$$

That is for Clifford frames the interplay between inner product $(\cdot, \cdot)$ *and Clifford algebra valued product* $\langle\langle \cdot, \cdot \rangle\rangle$ *on* H_n *is essential.*

EXAMPLE 5.2.1 (A Clifford Frame for $\mathbb{H}$):
$\{i\}$ is a tight Clifford frame for the quaternions. Let $0 \neq f = f_r + i f_i + j f_j + k f_k \in \mathbb{H}$. Then $|\langle\langle f, i\rangle\rangle| = |i f_r + f_i + k f_j - i f_k| = |f|$. A corresponding dual frame is $\{-i\}$. Note that $\{i\}$ is not a frame for the real Hilbert space $\big(\mathbb{H}, (\cdot,\cdot)\big)$.

EXAMPLE 5.2.2:
LET $\{f_k\}_{k\in\mathbb{N}}$ be a frame for a Hilbert space H.
THEN $\{e_0 f_k\}_{k\in\mathbb{N}}$ is a Clifford frame for H_n.

Proof. To verify that $\{e_0 f_k\}_{k\in\mathbb{N}}$ is a Clifford frame, let $g = \sum_{\boldsymbol{\alpha}\in\mathscr{O}_n} e_{\boldsymbol{\alpha}} g_{\boldsymbol{\alpha}} \in H_n$. Then

$$\begin{aligned}
\sum_{k\in\mathbb{N}} \left\| \langle\langle g, e_0 f_k \rangle\rangle \right\|^2 &= \sum_{k\in\mathbb{N}} \left\| \langle\langle \sum_{\boldsymbol{\alpha}\in\mathscr{O}_n} e_{\boldsymbol{\alpha}} g_{\boldsymbol{\alpha}}, e_0 f_k \rangle\rangle \right\|^2 \\
&= \sum_{k} \left\| \sum_{\boldsymbol{\alpha}\in\mathscr{O}_n} e_{\boldsymbol{\alpha}} \langle g_{\boldsymbol{\alpha}}, f_k\rangle \right\|^2 = \sum_{\boldsymbol{\alpha}\in\mathscr{O}_n} \sum_{k} \left| \langle g_{\boldsymbol{\alpha}}, f_k \rangle \right|^2 \\
&\leq B \sum_{\boldsymbol{\alpha}\in\mathscr{O}_n} \|g_{\boldsymbol{\alpha}}\|_H^2 \\
&= B \|g\|_{H_n}^2 .
\end{aligned}$$

A lower frame bound can be derived in a similar way. □

Lemma 5.2.9 (Properties of the frame operator and the canonical dual frame)
LET $\{f_k\}_{k\in\mathbb{N}} \subset H_n$ *be a Clifford frame with frame operator* S *and frame bounds* A, B.
THEN

(i) The frame operator S *is bounded, self-adjoint, invertible and positive. (An operator is called* **positive**, *iff* $\exists 0 < A : A\|f\|^2 \leq (Sf, f)$.*)*

(ii) $\{S^{-1} f_k\}_{k\in\mathbb{N}}$ *is a Clifford frame, called the* **canonical dual frame**, *with frame bounds* B^{-1}, A^{-1}. *The frame operator for* $\{S^{-1} f_k\}_{k\in\mathbb{N}}$ *is* S^{-1}.

Proof. ad(i) S is bounded as it is the composition of bounded operators. By Theorem 5.2.3

$$\|S\| = \|TT^*\| = \|T\|\,\|T^*\| = \|T\|^2 \leq B.$$

Since $S^* = (TT^*)^* = TT^* = S$ the frame operator is self-adjoint.

Now

$$\begin{aligned}
(Sf, f) = \Big(\sum_{k\in\mathbb{N}} \langle\langle f, f_k \rangle\rangle f_k, f \Big) &= \Big\langle \langle\langle \sum_{k\in\mathbb{N}} \langle\langle f, f_k\rangle\rangle f_k, f \rangle\rangle \Big\rangle_0 \\
&\overset{3.2.4(i)}{=} \sum_{k\in\mathbb{N}} \Big\langle \langle\langle f, f_k\rangle\rangle \langle\langle f_k, f\rangle\rangle \Big\rangle_0 \\
&\overset{3.2.4(iv)}{=} \sum_{k\in\mathbb{N}} |\langle\langle f, f_k\rangle\rangle|^2
\end{aligned}$$

From (5.2) we know that (Sf, f) is positive and bounded for all $f \in H_n$. We write

$$A\,\mathrm{Id} \leq S \leq B\,\mathrm{Id}. \tag{5.4}$$

Since S is positive, we know that $0 \leq \mathrm{Id} - B^{-1}S \leq \frac{B-A}{B}\,\mathrm{Id} < \mathrm{Id}$. That is,

$$\|\,\mathrm{Id} - B^{-1}S\| = \sup_{\|f\|\leq 1} \left|\big((\mathrm{Id} - B^{-1}S)f, f\big)\right| \leq \frac{B-A}{B} < 1.$$

Hence S is invertible by its Neumann series. (See [29] chapter I, example 4.5.)

Furthermore, S^{-1} is computed using a series of positiv $\mathbb{C}_n$-left-linear operators, and consequently S^{-1} is positive and $\mathbb{C}_n$-left-linear.

ad(ii)

$$\begin{aligned}\sum_{k\in\mathbb{N}} \left|\langle\!\langle f, S^{-1}f_k\rangle\!\rangle\right|^2 &= \sum_{k\in\mathbb{N}} \left|\langle\!\langle S^{-1}f, f_k\rangle\!\rangle\right|^2 \leq B\|S^{-1}f\|^2 \\ &\leq B\|S^{-1}\|^2\|f\|^2, \quad \forall f \in H_n.\end{aligned}$$

Thus $\{S^{-1}f_k\}_{k\in\mathbb{N}}$ is a Bessel sequence. That is, its frame operator is well defined. Now, since S is $\mathbb{C}_n$-left-linear, so is S^{-1} and hence

$$\sum_{k\in\mathbb{N}} \langle\!\langle f, S^{-1}f_k\rangle\!\rangle S^{-1}f_k = S^{-1}\sum_{k\in\mathbb{N}} \langle\!\langle S^{-1}f, f_k\rangle\!\rangle f_k = S^{-1}SS^{-1}f = S^{-1}f. \tag{5.5}$$

That is, the frame Operator of $\{S^{-1}f_k\}_{k\in\mathbb{N}}$ is S^{-1}.

To compute the frame bounds for $\{S^{-1}f_k\}_{k\in\mathbb{N}}$, we multiply (5.4) by S^{-1} to obtain

$$AS^{-1} \leq \mathrm{Id} \leq BS^{-1}.$$

That is,

$$B^{-1}\|f\|^2 \leq (S^{-1}f, f) \leq A^{-1}\|f\|^2, \quad \forall f \in H_n.$$

Since S^{-1} is the frame operator of $\{S^{-1}f_k\}_{k\in\mathbb{N}}$, it follows by (5.5) that

$$B^{-1}\|f\|^2 \leq \sum_{k\in\mathbb{N}} \left|\langle\!\langle f, S^{-1}f_k\rangle\!\rangle\right|^2 \leq A^{-1}\|f\|^2, \quad \forall f \in H_n.$$

□

The next theorem shows that the canonical dual frame defined in Lemma 5.2.9 is indeed a dual frame according to the definition in Definition 5.2.7.

Theorem 5.2.10 (Existence of a dual frame)
LET $\{f_k\}_{k\in\mathbb{N}} \subset H$ *be a Clifford frame with frame operator S. Then*

$$f = \sum_{k\in\mathbb{N}} \langle\!\langle f, S^{-1}f_k\rangle\!\rangle f_k, \quad \forall f \in H_n,$$

and

$$f = \sum_{k\in\mathbb{N}} \langle\!\langle f, f_k\rangle\!\rangle S^{-1}f_k, \quad \forall f \in H_n.$$

Both series converge unconditionally for all $f \in H_n$.

Proof. Let $f \in H_n$. On the one hand

$$f = SS^{-1}f = \sum_{k\in\mathbb{N}} \langle\!\langle S^{-1}f, f_k\rangle\!\rangle f_k = \sum_{k\in\mathbb{N}} \langle\!\langle f, S^{-1}f_k\rangle\!\rangle f_k,$$

and on the other hand

$$f = S^{-1}Sf = S^{-1}\sum_{k\in\mathbb{N}} \langle\langle f, f_k\rangle\rangle f_k = \sum_{k\in\mathbb{N}} \langle\langle f, f_k\rangle\rangle S^{-1} f_k.$$

The unconditional convergence follows since $\{f_k\}_{k\in\mathbb{N}}$ and $\{S^{-1}f_k\}_{k\in\mathbb{N}}$ are Bessel sequences and $T^* f \in l_n^2(\mathbb{N})$. □

Remark 5.2.11 (Existence of a frame decomposition)
The central property of Clifford frames is that they allow a basis-like decomposition – the frame decomposition – of a function in a Clifford-Hilbert module into a series in $l_n^2(\mathbb{N})$ by the analysis operator and reconstruction from this series via the synthesis operator of a dual frame. Theorem 5.2.10 shows that at least one dual frame exists – the canonical dual frame.

Lemma 5.2.12 (Frames on dense sets)
LET $\{f_k\}_{k\in\mathbb{N}} \subset H$. *If there exist* $0 < A \le B < \infty$ *such that for a dense subset* $V \subseteq H_n$

$$A\|f\|^2 \le \sum_{k\in\mathbb{N}} \left|\langle\langle f, f_k\rangle\rangle\right|^2 \le B\|f\|^2, \quad \forall f \in V.$$

THEN $\{f_k\}_{k\in\mathbb{N}}$ *is a Clifford frame for* H_n *with frame bounds* A, B.

Proof. In Lemma 5.2.6 we showed that the upper bound holds. Now we know by Lemma 5.2.9(ii) that for the dual frame $\{S^{-1}f_k\}$ it is true that

$$\sum_{k\in\mathbb{N}} \left|\langle\langle f, S^{-1}f_k\rangle\rangle\right|^2 \le \frac{1}{A}\|f\|^2 \quad \forall f \in V.$$

Hence by Lemma 5.2.6

$$\sum_{k\in\mathbb{N}} \left|\langle\langle f, S^{-1}f_k\rangle\rangle\right|^2 \le \frac{1}{A}\|f\|^2 \quad \forall f \in \mathbb{H}_n,$$

and again by Lemma 5.2.9(ii) the lower bound follows. □

Theorem 1.3.27 on page 13 shows that a surjective operator maps a frame to a frame of its range. The next theorem is the analog of Theorem 1.3.27 for Clifford frames. In Example 5.1.2 we used Theorem 1.3.27 to prove that the Hilbert transform of a frame is once again a frame. The next theorem will play this role for the Riesz transform of a frame.

Theorem 5.2.13 (Clifford frames and operators)
LET $\{f_k\}_{k\in\mathbb{N}} \subset H_n$ *be a Clifford frame for* H_n *with bounds* A, B. *Let* $V \subseteq H_n$. *Furthermore, let* $U : H_n \mapsto V$ *be a bounded surjective operator.*
THEN $\{Uf_k\}_{k\in\mathbb{N}}$ *is a Clifford frame for* $\overline{V}$ *with bounds* $A\|U^\dagger\|^{-2}, B\|U\|^2$. *Here* $U^\dagger = U^*(UU^*)^{-1}$ *denotes the pseudoinverse.*

Proof. Let $f \in V$. Then

$$\sum_{k\in\mathbb{N}} \left|\langle\langle f, Uf_k\rangle\rangle\right|^2 \le B\|U^* f\|^2 \le B\|U\|^2\|f\|^2,$$

whence $\{Uf_k\}_{k\in\mathbb{N}}$ is a Bessel sequence.
Since U is surjective, there exists a $g \in H_n$ such that $f = Ug$.

Let $U^\dagger := U^*(UU^*)^{-1}$ be the pseudo-inverse of U. For details on the pseudo-inverse of an operator on a Hilbert space see [10] A.7.
Now $UU^\dagger = UU^*(UU^*)^{-1} = \mathrm{Id}_{H_n}$ is self adjoint. Therefore

$$f = Ug = (UU^\dagger)^* Ug = (U^\dagger)^* U^* Ug.$$

And hence

$$\begin{aligned}\|f\|^2 \le \|(U^\dagger)^*\|^2 \|U^* Ug\|^2 \le A^{-1}\|(U^\dagger)^*\|^2 \sum_{k\in\mathbb{N}} \left|\langle\!\langle U^* Ug, f_k\rangle\!\rangle\right|^2 \\ = A^{-1}\|U^\dagger\|^2 \sum_{k\in\mathbb{N}} \left|\langle\!\langle f, Uf_k\rangle\!\rangle\right|^2.\end{aligned}$$

Now Lemma 5.2.12 yields the frame property on $\overline{V}$. □

5.3 Monogenic wavelets

In the last section we have made the first step towards extending the analytical wavelets introduced in Example 5.1.2 to higher dimensions. In this section we will introduce a suitable wavelet transform and finish our agenda to define monogenic wavelet frames by Definition 5.3.2.

Definition 5.3.1 (Hypercomplex wavelet transform)
LET $\psi \in L^2(\mathbb{R}^n)_n$ *and let* D *be a dilation matrix.*
THEN

$$\{D^j T_k \psi\}_{j,k}$$

is called a **wavelet system**, *where* $j \in \mathbb{Z}$ *and* $k \in \mathbb{Z}^n$ *or* $j \in \mathbb{R}$ *and* $k \in \mathbb{R}^n$ *with* **mother wavelet** ψ.
LET $f \in L^2(\mathbb{R}^n)_n$.
THEN

$$W_\psi(f) := \left\{\langle\!\langle f, D^j T_k \psi\rangle\!\rangle\right\}_{j,k}$$

is called the **hypercomplex wavelet transform** *of* f.

Definition 5.3.2 (Monogenic wavelet)
LET $\psi \in L^2(\mathbb{R}^n)$ *be a mother wavelet for* $L^2(\mathbb{R}^n)$ *with respect to a dilation matrix* D.
THEN *the* **monogenic wavelet system** $\{D^j T_k \psi_m\}_{j,k}$ *corresponding to* ψ *is generated by the* **monogenic mother wavelet**

$$\psi_m := \psi + R\psi = \psi + \sum e_\alpha R_\alpha \psi,$$

where R *is the hypercomplex Riesz transformation defined in Definition 4.1.1 on page 66. An element* $D^j T_k \psi$ *of the monogenic wavelet system is called a* **monogenic wavelet**. *The hypercomplex wavelet transform* W_{ψ_m} *is called the* **monogenic wavelet transform**.

5.3.1 Monogenic wavelet frames

Theorem 5.3.3 (Riesz transforms of frames)
LET $\{f_k\}_{k\in\mathbb{N}}$ *be a frame for* $L^2(\mathbb{R}^n,\mathbb{R})$ *with frame bounds* A *and* B.
THEN *the following hold:*

(i) $\{f_k\}_{k\in\mathbb{N}}$ *is a Clifford frame for* $L^2(\mathbb{R}^n,\mathbb{C})_n \cong L^2(\mathbb{R}^n,\mathbb{C}_n)$ *with the same frame bounds* A *and* B.

(ii) The Riesz transformed frame $\{Rf_k\}_{k\in\mathbb{N}}$ *is a Clifford frame for* $L^2(\mathbb{R}^n,\mathbb{C})_n$ *with the same frame bounds* A *and* B.

Proof. Let $\{f_k\}_{k\in\mathbb{N}}$ be a frame for $L^2(\mathbb{R}^n,\mathbb{R})$. Then $\{e_0 f_k\}_{k\in\mathbb{N}}$ is a frame for $L^2(\mathbb{R}^n,\mathbb{C}_n)$. To verify this, let $g=\sum_{\boldsymbol{\alpha}\in\mathcal{O}_n} e_{\boldsymbol{\alpha}} g_{\boldsymbol{\alpha}} \in L^2(\mathbb{R}^n,\mathbb{C}_n)$. Thus

$$\begin{aligned}
\sum_{k\in\mathbb{N}} \left\| \langle\langle g, e_0 f_k\rangle\rangle \right\|^2 &= \sum_{k\in\mathbb{N}} \left\| \langle\langle \sum_{\boldsymbol{\alpha}\in\mathcal{O}_n} e_{\boldsymbol{\alpha}} g_{\boldsymbol{\alpha}}, e_0 f_k\rangle\rangle \right\|^2 \\
&= \sum_k \left\| \sum_{\boldsymbol{\alpha}\in\mathcal{O}_n} e_{\boldsymbol{\alpha}} \langle g_{\boldsymbol{\alpha}}, f_k\rangle \right\|^2 = \sum_{\boldsymbol{\alpha}\in\mathcal{O}_n} \sum_k \left| \langle g_{\boldsymbol{\alpha}}, f_k\rangle \right|^2 \\
&\le B \sum_{\boldsymbol{\alpha}\in\mathcal{O}_n} \|g_{\boldsymbol{\alpha}}\|^2_{L^2(\mathbb{R}^n,\mathbb{C})} \\
&= B\|g\|^2_{L^2(\mathbb{R}^n,\mathbb{C})_n}.
\end{aligned}$$

A lower frame bound can be derived in a similar way. This proves (i).

The Riesz transform is its own inverse, hence invertible and therefore surjective in $L^2(\mathbb{R}_n,\mathbb{C}_n)$. (See Theorem 4.1.4 on page 66.) Now statement (ii) follows by Theorem 5.2.13.

Summing up the Riesz transform of a frame for $L^2(\mathbb{R}^n,\mathbb{R})$ yields a Clifford frame for $L^2(\mathbb{R}^n,\mathbb{C}_n)$. □

Theorem 5.3.4 (Monogenic wavelet frame)
LET $\psi \in L^2(\mathbb{R}^n)$ *generate a wavelet frame for* $L^2(\mathbb{R}^n)$ *with respect to a dilation matrix* D *with frame bounds* $0 < m \le M < \infty$.

THEN *the monogenic wavelet generates a wavelet frame for* $TH^2_-(\mathbb{R}^n)$ *with frame bounds* $2m$ *and* $2M$.

LET $\phi \in L^2(\mathbb{R}^n)$ *generate a dual frame for the frame with respect to the same dilation and translation.*

THEN *the dual frame of the monogenic wavelet frame with mother wavelet* ψ_m *is the frame with mother wavelet* $\frac{1}{2}\phi_m$.

LET *the dilation matrix constitute a rotated dilation, i.e.* $D=\mathfrak{D}_d := D_d\rho$, *where* $d>1$ *and* $\rho\in SO(n)$.

THEN *the wavelet transform of a function in* $L^2(\mathbb{R}^n)$ *satisfies*

$$W_{\psi_m} f(t,j) = W_\psi f(t,j) + \sum_{l,k} e_l (\rho^j)_{k,l} W_\psi R_k f(t,j).$$

Thus the monogenic signal of a function $f\in L^2(\mathbb{R}^n)$ *may be computed from the coefficients of the monogenic wavelet.*

Proof. Since ψ generates a frame for $L^2(\mathbb{R}^n)$ with real frame coefficients, using Clifford algebra valued frame coefficients it generates a frame for $L^2_n(\mathbb{R}^n)$. $TH^2_-(\mathbb{R}^n)$ is the image of $e_o L^2(\mathbb{R}^n)$ under the monogenic signal and by Theorem 5.3.3 ψ_m is a mother wavelet for a wavelet frame of $TH^2_-(\mathbb{R}^n)$.

Let $f \in L^2_n(\mathbb{R}^n)$, $t \in \mathbb{Z}^n$ and $j \in \mathbb{Z}$.

$$\begin{aligned}
W_{\psi_m} f(t,j) &= \langle\!\langle f, \mathfrak{D}_d^j T_t \psi_m \rangle\!\rangle \\
&= \langle\!\langle f, \mathfrak{D}_d^j T_t (\psi + R\psi) \rangle\!\rangle \\
&= \langle f, \mathfrak{D}_d^j T_t \psi \rangle + \sum_{l=1}^n e_l i \langle\!\langle f, D_{d^j} \rho^j T_t R_l \psi \rangle\!\rangle \\
&= \langle f, \mathfrak{D}_d^j T_t \psi \rangle + \sum_{l,k=1}^n (\rho^j)_{k,l} e_l i \langle f, R_k D_{d^j} \rho^j T_t \psi \rangle \\
&= \langle f, \mathfrak{D}_d^j T_t \psi \rangle + \sum_{l,k=1}^n (\rho^j)_{k,l} \int_{\mathbb{R}^n} e_l \widehat{f}(\xi) \overline{i \frac{\xi_k}{|\xi|} \mathscr{F}\big(D_{d^j} \rho^j T_t \psi\big)(\xi)} d\xi \\
&= \langle f, \mathfrak{D}_d^j T_t \psi \rangle + \sum_{l,k=1}^n (\rho^j)_{k,l} \int_{\mathbb{R}^n} e_l i \widehat{f}(\xi) \overline{\mathscr{F}\big(\mathfrak{D}_d^j T_t \psi\big)(\xi) \frac{\xi_k}{|\xi|}} d\xi \\
&= \langle f, \mathfrak{D}_d^j T_t \psi \rangle + \sum_{l,k=1}^n (\rho^j)_{k,l} e_l \langle R_k f, \mathfrak{D}_d^j T_t \psi \rangle.
\end{aligned}$$

Here we used Theorem 2.1.2 (i, ii, iii, vii) on page 20.

We now show that $\forall f, g \in L^2(\mathbb{R}^n)$ the following identity holds

$$\Big(\sum_{\substack{j \in \mathbb{Z}, \\ t \in \mathbb{Z}^n}} \langle\!\langle e_0 f, \mathfrak{D}_d^j T_t \psi_m \rangle\!\rangle \mathfrak{D}_d^j T_t \phi_m, e_0 g \Big) = 2 \langle f, g \rangle,$$

using what we have just proven.

$$\begin{aligned}
\Big(\sum_{\substack{j \in \mathbb{Z}, \\ t \in \mathbb{Z}^n}} \langle\!\langle e_0 f, \mathfrak{D}_d^j T_t \psi_m \rangle\!\rangle \mathfrak{D}_d^j T_t \phi_m, e_0 g \Big) &= \sum_{j,t} \Big\langle \langle\!\langle e_0 f, \mathfrak{D}_d^j T_t \psi_m \rangle\!\rangle \langle\!\langle \mathfrak{D}_d^j T_t \phi_m, e_0 g \rangle\!\rangle \Big\rangle_0 \\
&= \sum_{j,t} \langle f, \mathfrak{D}_d^j T_t \psi \rangle \langle \mathfrak{D}_d^j T_t \phi, g \rangle + \Big\langle \sum_{k,l=1}^n (\rho^j)_{k,l} e_l \langle R_k f, \mathfrak{D}_d^j T_t \psi \rangle \sum_{r,s=1}^n (\rho^j)_{r,s} \overline{e_s} \langle \mathfrak{D}_d^j T_t \phi, R_r g \rangle \Big\rangle_0 \\
&= \Big\langle \sum_{j,t} \langle f, \mathfrak{D}_d^j T_t \psi \rangle \mathfrak{D}_d^j T_t \phi, g \Big\rangle + \sum_{j,t} \sum_{k,r,l=1}^n \underbrace{(\rho^j)_{k,l} (\rho^j)_{r,l}}_{=(\rho\rho^T)_{k,r} = \mathrm{Id}_{k,r} = \delta_{k,r}} \langle R_k f, \mathfrak{D}_d^j T_t \psi \rangle \langle \mathfrak{D}_d^j T_t \phi, R_r g \rangle \\
&= \langle f, g \rangle + \sum_{k=1}^n \sum_{j,t} \langle R_k f, \mathfrak{D}_d^j T_t \psi \rangle \langle \mathfrak{D}_d^j T_t \phi, R_k g \rangle \\
&= \langle f, g \rangle + \sum_{k=1}^n \Big\langle \sum_{j,t} \langle R_k f, \mathfrak{D}_d^j T_t \psi \rangle \mathfrak{D}_d^j T_t \phi, R_k g \Big\rangle \\
&= \langle f, g \rangle + \sum_{k=1}^n \langle R_k f, R_k g \rangle \\
&= 2 \langle f, g \rangle, \ \forall f, g \in L^2(\mathbb{R}^n).
\end{aligned}$$

□

5.4 Decay of monogenic wavelets

The joint localization of wavelets in space and frequency is one of their key properties. As a first step we now look at the decay rate of Riesz transformed functions. The Riesz transform leaves the decay in the Fourier domain invariant. In the following we derive novel conditions under which a partial Riesz transform of a function inherits the decay rate.

We take a closer look at the joint space-frequency localization of the monogenic wavelets when we consider the effect of the Riesz transform on uncertainty relations in Theorem 6.3.12.

Lemma 5.4.1 (Smoothness and decay)
LET $k \in \mathbb{N}_0$ *and* $f \in \mathscr{S}'(\mathbb{R}^n)$.
THEN

$$\partial^{\boldsymbol{\beta}} f \in L^2(\mathbb{R}^n),\ \forall |\boldsymbol{\beta}| \le k \text{ iff } \int_{\mathbb{R}^n} |\widehat{f}(\xi)|^2 (1+|\xi|^2)^k d\xi < \infty.$$

Here the derivative operator should be considered in the distributional sense.

Proof. The proof can be found in [20], Chapter 1.2. □

Proposition 5.4.2
LET $\widehat{f} \in C^k(\mathbb{R}^n)$ *and furthermore let* $\widehat{f}^{(\boldsymbol{\alpha})}(0) = 0,\ \forall |\boldsymbol{\alpha}| \le k$.
THEN $\widehat{R_\alpha f} \in C^k(\mathbb{R}^n)$.

Proof. Remember from Theorem 2.2.8 on page 30 that the partial derivatives of the Fourier multiplier of the Riesz transform are of the form

$$\partial^{\boldsymbol{\beta}} \frac{\xi_\alpha}{|\xi|} = \frac{h_{\alpha,\boldsymbol{\beta}}(\xi)}{|\xi|^{|\boldsymbol{\beta}|}},\ \forall \xi \in \mathbb{R}^n \setminus \{0\}, \tag{5.6}$$

where $h_{\alpha,\boldsymbol{\beta}} : \mathbb{R}^n \to \mathbb{R}$ is a 0-homogenous function.

To show that $\widehat{R_\alpha f} \in C^k(\mathbb{R}^n)$ we have to show that all partial derivatives $\partial^{\boldsymbol{\beta}} \widehat{R_\alpha f}$, $|\boldsymbol{\beta}| \le k$, are continuous.

Since $\partial^{\boldsymbol{\beta}} \widehat{R_\alpha} \in C^\infty(\mathbb{R}^n \setminus \overline{B_\epsilon(0)}),\ \forall \epsilon > 0$ it is sufficient to show that

$$\lim_{x \to 0} \partial^{\boldsymbol{\beta}} \widehat{R_\alpha f}(\xi) = 0,\ \forall |\boldsymbol{\beta}| \le k.$$

For $\xi \in \mathbb{R}^n$, the product rule yields

$$\partial^{\boldsymbol{\beta}} \widehat{R_\alpha f}(\xi) = \sum_{\gamma \le \beta}^{n} \binom{n}{\gamma_1} \cdots \binom{n}{\gamma_n} \partial^\gamma \frac{\xi_\alpha}{|\xi|} \partial^{\boldsymbol{\beta}-\gamma} \widehat{f}(\xi),$$

where $\gamma = (\gamma_1, \cdots, \gamma_n) \in \mathbb{N}_0^n$.

Combined with Equation 5.6 this yields a term of the form

$$\partial^{\boldsymbol{\beta}} \widehat{R_\alpha f}(\xi) = \sum_{\gamma \le \boldsymbol{\beta}} \binom{n}{\gamma_1} \cdots \binom{n}{\gamma_n} \frac{h_{\alpha,\gamma}(\xi)}{|\xi|^{|\gamma|}} \partial^{\boldsymbol{\beta}-\gamma} \widehat{f}(\xi),\ \forall \xi \in \mathbb{R}^n \setminus \{0\}.$$

By Taylors Theorem A.1.3 on page 166 and since $\partial^{\boldsymbol{\nu}} \widehat{f}(0) = 0,\ \forall |\boldsymbol{\nu}| \le k$, we know that

$$0 = \lim_{\xi \to 0} \frac{\partial^{\boldsymbol{\beta}-\gamma} \widehat{f}(\xi) - T_{|\gamma|} \partial^{\boldsymbol{\beta}-\gamma} \widehat{f}(\xi; 0)}{|\xi|^{|\gamma|}} = \lim_{\xi \to 0} \frac{\partial^{\boldsymbol{\beta}-\gamma} \widehat{f}(\xi)}{|\xi|^{|\gamma|}},$$

where $T_{|\gamma|}\partial^{\boldsymbol{\beta}-\gamma}\widehat{f}(\xi;a) = \sum_{l=0}^{|\gamma|}\frac{1}{l!}\sum_{\substack{\boldsymbol{\alpha}\in\mathbb{N}_0^n\\|\boldsymbol{\alpha}|=l}}\partial^{\boldsymbol{\alpha}+\boldsymbol{\beta}-\gamma}f(0)\xi^{\boldsymbol{\alpha}}$.

□

Theorem 5.4.3 (Vanishing moments and decay of the Riesz transform)
LET $k \in \mathbb{N}_0$. *Furthermore let* $f \in L^2(\mathbb{R}^n)$ *be such that* $\partial^{\boldsymbol{\beta}}\widehat{f} \in L^2(\mathbb{R}^n)$, $\forall|\boldsymbol{\beta}| \le k$ *and* $\widehat{f} \in C^k$. *Finally, let* $\partial^{\boldsymbol{\beta}} f(0) = 0$, $\forall|\boldsymbol{\beta}| \le k$.
THEN $\widehat{R_\alpha f} \in W^{k,2}(\mathbb{R}^n)$. *I.e.*

$$\int_{\mathbb{R}^n}|R_\alpha f(\xi)|^2(1+|\xi|^2)^k d\xi < \infty.$$

Proof. We want to use Lemma 5.4.1. Hence we need to proof that

$$\partial^{\boldsymbol{\beta}}\widehat{R_\alpha f} \in L^2(\mathbb{R}^n),\ \forall|\boldsymbol{\beta}| \le k.$$

In Proposition 5.4.2 we proved the existence of the derivatives. Square-integrability follows easily if we consider the integrability on $\overline{B_1(0)}$ and on $\mathbb{R}^n \setminus \overline{B_1(0)}$ separately. We have already shown, that $\partial^{\boldsymbol{\beta}}\widehat{R_\alpha f}$ is continuous on the compact set $\overline{B_1(0)}$, hence boundedness and square-integrability follows. As we have seen in Remark 2.2.7 on page 29

$$\partial^{\boldsymbol{\beta}}\widehat{R_\alpha f}(\xi) = \sum_{\gamma\le\beta}\binom{n}{\gamma_1}\cdots\binom{n}{\gamma_n}\partial^\gamma\widehat{f}(\xi)\frac{p_{\alpha,\gamma}(\xi)}{q_{2|\gamma|+1}(|\xi|)},\ \forall\xi\in\mathbb{R}^n.$$

But $\partial^\gamma\widehat{f} \in L^2(\mathbb{R}^n)$ and $\frac{p_{\alpha,\gamma}(\xi)}{q_{2|\gamma|+1}(|\xi|)}$ is dominated by a finite constant on $(\mathbb{R}^n)\setminus\overline{B_1(0)}$. Hence square-integrability follows. □

Remark 5.4.4
An important condition of Theorem 5.4.3 is the existence of the derivatives of $\widehat{f}$ *of order less or equal to* k. *The existence of the derivatives, at least in a neighborhood of* 0, *is indeed necessary for our proof. It would be sufficient to have weak derivatives of degree* k *on the rest of* $\mathbb{R}^n$.

5.5 Explicit construction of wavelet frames in $L^2(\mathbb{R}^n)$

On our way to implement the monogenic signal via wavelet frames we have found in Theorem 5.3.4 that we can derive a monogenic wavelet frame from a real wavelet frame of $L^2(\mathbb{R}^n)$. Hence our next step should be to construct suitable wavelet frames for $L^2(\mathbb{R}^n)$.

The present section presents a suitable way to construct wavelet frames for $L^2(\mathbb{R}^n)$ with respect to a given isotropic or rotated dilation.

5.5.1 Wavelets on irregular grids with arbitrary dilation matrices, and frame atoms for $L^2(\mathbb{R}^n)$

This section introduces a construction principle for frames of $L^2(\mathbb{R}^n)$ based on [2]. We finish the section with our own Corollary 5.5.6 that shows how to acquire tight wavelet frames.

Proposition 5.5.1
LET U, V *be non-empty, open subsets of* $\mathbb{R}^n$, *and let* $T : U \mapsto V$ *be a diffeomorphism, i.e.,*

$T \in C^1(\mathbb{R}^n)$ and $T^{-1} \in C^1(\mathbb{R}^n)$, such that

$$0 < \alpha := \inf_{y \in U} |\det(T'(y)|^{-1} \leq \beta := \sup_{y \in U} |\det(T'(y))|^{-1} < \infty,$$

where $T'(y) := \left(\frac{\partial T_i}{\partial y_j}(y)\right)_{i,j}$. Furthermore, let $\{g_j\}_{j\in\mathbb{Z}}$ be a frame for

$$K_V := \left\{f \in L^2(\mathbb{R}^n) : \operatorname{supp}(f) \subseteq V\right\}$$

with lower frame bound m and and upper frame bound M such that $0 < m < M < \infty$.

THEN *$\{g_j \circ T\}_{j\in\mathbb{Z}}$ is a frame for $K_U := \left\{f \in L^2(\mathbb{R}^n) : \operatorname{supp}(f) \subseteq U\right\}$ with lower frame bound αm and upper frame bound βM.*

Proof. Let $f \in K_U$. We will write $S := T^{-1}$, with this notation

$$\begin{aligned} \langle f, g_j \circ T\rangle &= \int_U f(t)\overline{g_j(T(t))}dt = \int_V f(S(s))|\det(S'(s))|\overline{g_j(s)}ds \\ &= \langle |\det(S'(.))| f \circ S, g_j\rangle \end{aligned}$$

Now, $|\det(S'(.))|f \circ S \in K_V$, since $0 < |\det(S'(s))| < \infty$, $\forall s \in V$. It follows that

$$m \| |\det(S'(.))| f \circ S\|_2^2 \leq \sum_{j\in\mathbb{Z}} |\langle |\det(S'(.))| f \circ S, g_j|^2 \leq M \| |\det(S'(.))| f \circ S\|_2^2.$$

Furthermore,

$$\begin{aligned} \| |\det(S'(.))| f \circ S\|_2^2 &= \int_V |\det(S'(t))|^2 |f(S(t))|^2 dt \\ &= \int_U |\det(S' \circ T(s))|^2 |\det(T'(s))| |f(s)|^2 ds. \end{aligned}$$

Now since $|\det(S' \circ T(.))| = \left(|\det(T'(.))|\right)^{-1}$ and $\alpha\|f\|^2 \leq \int_U |\det(T'(.))|^{-1}|f(t)|^2 \leq \beta\|f\|^2 dt$ it follows that

$$m\alpha\|f\|^2 \leq \sum_{j\in\mathbb{Z}} |\langle f, g_j \circ T\rangle|^2 \leq \beta M \|f\|^2.$$

□

Corollary 5.5.2
LET *$Q \subseteq \mathbb{R}^n$ be an open subset, $y \in \mathbb{R}^n$ and let $A \in GL_n(\mathbb{R}^n)$ be invertible.*

THEN *$\{g_j\}_{j\in\mathbb{Z}}$ is a frame for K_Q, with frame bounds m, M, iff*

(i) $\{T_y g_j\}_{j\in\mathbb{Z}}$ is a frame for $K_{T_{-y}Q}$ with frame bounds m, M.

(ii) $\{A g_j\}_{j\in\mathbb{Z}}$ is a frame for $K_{A^{-1}Q}$ with frame bounds m, M.

Definition 5.5.3 (Riesz partition of unity)
LET *$\{h_j\}_{j\in\mathbb{Z}} \subset L^2(\mathbb{R}^n)$ be a set of functions in $L^2(\mathbb{R}^n)$ such that there exist positive finite numbers $0 < p \leq P < \infty$ such that*

$$p \leq \sum_{j\in\mathbb{Z}} |h_j(x)|^2 \leq P, \text{ a.e. } x \in \mathbb{R}^n. \tag{5.7}$$

THEN *$\{h_j\}_{j\in\mathbb{Z}} \subset L^2(\mathbb{R}^n)$ is a* **Riesz partition of unity**.

LET $p = P = 1$.

THEN *$H = \{h_j\}_{j\in\mathbb{Z}}$ is called a* **regular partition of unity**.

The next theorem reduces the problem to construct a frame for $L^2(\mathbb{R}^n)$ to the problem of finding a Riesz partition of unity.

Theorem 5.5.4 (Frames and Riesz partitions of unity)
LET $H = \{\overline{h_j}\}_{j\in\mathbb{Z}} \subset L^2(\mathbb{R}^n)$ *be a Riesz partition of unity with bounds* p, P *and denote*

$$W_j := \{\overline{h}_j f : f \in L^2(\mathbb{R}^n)\}.$$

Furthermore let $\{g_{j,k}\}_k$ *be a frame for* W_j *with frame bounds* m_j, M_j *and let*

$$0 < m = \inf_j m_j \le \sup_j M_j = M < \infty.$$

THEN $\{h_j g_{j,k} : j, k \in \mathbb{Z}\}$ *is a frame for* $L^2(\mathbb{R}^n)$ *with frame bounds* pm *and* PM.

Proof. Let $f \in L^2(\mathbb{R}^n)$ and set $f_j := \overline{h_j} f \in W_j$. Then

$$\begin{aligned}
mp\|f\|^2 &= \inf_j m_j \int_{\mathbb{R}^n} p|f(t)|^2 dt \le \inf_j m_j \int_{\mathbb{R}^n} \sum_j |\overline{h_j(t)}|^2 |f(t)|^2 dt \\
&= \inf_j m_j \int_{\mathbb{R}^n} \sum_j |\overline{h_j(t)} f(t)|^2 dt \le \sum_j m_j \|f_j(t)\|^2 \\
&\le \sum_j \sum_k |\langle f, \overline{h_j} g_{j,k}\rangle| \le \sum_j M_j \|f_j(t)\|^2 \\
&\le \sup_j M_j \int_{\mathbb{R}^n} \sum_j |\overline{h_j(t)}|^2 |f(t)|^2 dt \le \sup_j M_j \int_{\mathbb{R}^n} P|f(t)|^2 dt \\
&= MP\|f\|^2.
\end{aligned}$$

□

Remark 5.5.5
At first view Theorem 5.5.4 does not seem to make the problem of finding a wavelet frame for $L^2(\mathbb{R}^n)$ *any easier. If however the Riesz partition of unity is chosen as dilates of a single compactly supported function* $h \in L^2(\mathbb{R}^n)$*, then the* $\{g_{j,k}\}_{k\in\mathbb{Z}}$ *may be chosen as dilates of a set of modulations and* $\{\widehat{h_j g_{j,k}}\}_{j,k\in\mathbb{Z}}$ *is a wavelet frame for* $L^2(\mathbb{R}^n)$.

Since we wish to construct wavelet frames we have to make sure that the dual frame is a wavelet frame, too. The easiest way to do this is to construct a tight wavelet frame. The Riesz partition of unity H and the frame $\{g_{j,k}\}_{k\in\mathbb{Z}}$ for the corresponding W_j can be chosen to yield a tight wavelet frame as follows.

Corollary 5.5.6 (Tight wavelet frames from regular partitions of unity)
LET $h \in L^2(\mathbb{R}^n)$ *and let* $H = \{h_j := h(A^j\cdot) : j \in \mathbb{Z}\}$ *be a regular partition of unity, where* A *is a dilation. Let*

$$W_j := \{\overline{h}_j f : f \in L^2(\mathbb{R}^n)\}.$$

Furthermore, let $w > 0$ *be such that* $\operatorname{supp}(h) \subseteq [-\frac{1}{2}w, \frac{1}{2}w]^n$. *Finally, let*

$$g_{j,k}(x) := A^j g_{0,k}(x) := \frac{1}{w^n} A^j e^{\frac{-2\pi}{w} i\langle x,k\rangle}.$$

THEN $\{g_{j,k}\}_k$ *is a tight frame for* W_j *with frame bound* 1. *Furthermore,* $\{h_j g_{j,k}\}_{j,k\in\mathbb{Z}}$ *is a tight frame for* $L^2(\mathbb{R}^n)$ *with frame bound* 1 *and its Fourier transform* $\{\mathscr{F}(h_j g_{j,k})\}_{j,k\in\mathbb{Z}}$ *is a wavelet tight frame for* $L^2(\mathbb{R}^n)$ *with frame bound* 1.

Proof. $\{e^{-2\pi i\langle x,k\rangle}\}_{k\in\mathbb{Z}}$ is an orthonormal basis for $L^2(-\frac{1}{2},\frac{1}{2})$ (see for example [41]) and hence a tight frame with bound 1 for every subset of $L^2(-\frac{1}{2},\frac{1}{2})$. Thus by Corollary 5.5.2 $\{g_{0,k}\}_{k\in\mathbb{Z}}$ is a tight frame with bound 1 for W_0. Again by Corollary 5.5.2 $\{g_{j,k}\}_{k\in\mathbb{Z}}$ is a tightframe with bound 1 for W_j. Now Theorem 5.5.4 proves that $\{h_j g_{j,k} : j,k\in\mathbb{Z}\}$ is a tight frame for $L^2(\mathbb{R}^n)$ with frame bound 1.

By Theorem 1.3.22 on page 12 we know that

$$\mathscr{F}(h_j g_{j,k}) = \mathscr{F}\big(h(A^j\cdot)\frac{1}{w^n}A^j e^{-2\pi i\langle\cdot,k\rangle/w}\big) = \frac{1}{w^n}\mathscr{F}(A^j M_{-k/w}h) = \frac{1}{w^n}(A^T)^{-j}T_{-k/w}\widehat{h},$$

whence $\{\mathscr{F}(h_j g_{j,k}) : j,k\in\mathbb{Z}\}$ is a wavelet tight frame for $L^2(\mathbb{R}^n)$ with frame bound 1. □

5.5.2 Construction of regular partitions of unity

Let $\mathbb{T} := [-\frac{1}{2},\frac{1}{2}]^n$ and $w\mathbb{T} := [-\frac{1}{2}w,\frac{1}{2}w]^n$. Corollary 5.5.6 states that, in order to construct the wavelet frames with respect to a dilation matrix A, it is sufficient to construct a regular partition of unity of the form $H = \{h(A^j\cdot)\}_{j\in\mathbb{Z}}$ such that $\operatorname{supp}(h)\subseteq w\mathbb{T}$, for some $w>0$.

The following new theorem reduces the problem of finding a regular partition of unity to that of finding a certain function which then allows to easily control differentiability.

Theorem 5.5.7 (Construction of regular partitions of unity)
For a closed set A denote by $\mathring{A} := \bigcup\{O\subset A:\ O \text{ is open}\}$ the interior and by $\partial A := A\setminus\mathring{A}$ the boundary of A.

LET $S\subseteq\mathbb{T}$ *be a closed nonempty set and let A be a dilation matrix on $\mathbb{R}^n$ such that*

$$\bigcup_{j\in\mathbb{Z}} A^j S = \mathbb{R}^n\setminus\{0\} \text{ and } \mathring{S}\cap\mathring{AS} = \emptyset.$$

We set $S_\circ := \partial S\cap\partial A^{-1}S$ and suppose that $\partial S = S_\circ\oplus AS_\circ$.
Let $f\in C^l(\mathbb{R}^n)$, $l\in\mathbb{N}_0$ be such that $0\le f(x)\le 1/4, \forall x\in S$, and furthermore let on the boundary

$$f(x)=0,\ f(Ax)=1/4\,,\partial^\alpha f(x)=\partial^\alpha f(Ax)=0,\ \forall\alpha\in\mathbb{N}_0^n: 1\le|\alpha|\le l,\ \forall x\in S_\circ.$$

THEN

$$h(x) = \begin{cases}\cos(2\pi f(Ax)), & \forall x\in A^{-1}S;\\ \sin(2\pi f(x)), & \forall x\in S;\\ 0, & \text{otherwise,}\end{cases}$$

defines a regular partition of unity with respect to the dilation A by

$$H = \{h(A^j\cdot)\}_{j\in\mathbb{Z}}.$$

Furthermore, $h_j = h(A^j\ \cdot\ )\in C^l(\mathbb{R}^n)$, $\forall j\in\mathbb{Z}$.

Proof. We first show that H is a regular partition of unity. Let $x\in\mathbb{R}^n$. Then $\exists k\in\mathbb{Z}: x\in A^{-k}S$. Thus,

$$\sum_{j\in\mathbb{Z}}|h_j(x)|^2 = |\cos(2\pi f(A^{k-1}(Ax)))|^2 + |\sin(2\pi f(A^k x))|^2 = 1.$$

Note that $\partial S = A^{-1}S_\circ + S_\circ$ and $\partial A^{-1}S = A^{-2}S_\circ + A^{-1}S_\circ$. Now, $\forall x \in S_\circ$: $h(A^{-2}x) = \cos(\pi/2) = 0$, and $h(x) = \sin(0) = 0$, and $h(A^{-1}x) = \cos(0) = \sin(\pi/2) = 1$. Thus h is well defined and continuous.

To show that $h \in C^l(\mathbb{R}^n)$, we need to show that $\partial^\alpha h$ is a continuous function for all $\alpha \in \mathbb{N}_0^n$ such that $|\alpha| \le l$, $x \in A^{-1}S + S$, especially in ∂S and $\partial A^{-1}S$.

$\partial^\alpha h$ is of the form

$$\partial^\alpha h(x) = \begin{cases} \sum_{k=1}^{l} \sum_{\substack{\beta\le\alpha \\ |\beta|=k}} c_\beta \partial^{\alpha-\beta}\Big(\prod_{m=1}^{n}(1-\delta_{\beta_m,0})\dfrac{\partial^{\beta_m}(2\pi f(Ax))}{\partial x_m}\Big)\cos^{(k)}(2\pi f(Ax)), & \forall x \in A^{-1}S, \\ \sum_{k=1}^{l} \sum_{\substack{\beta\le\alpha \\ |\beta|=k}} c_\beta \partial^{\alpha-\beta}\Big(\prod_{m=1}^{n}(1-\delta_{\beta_m,0})\dfrac{\partial^{\beta_m}(2\pi f(x))}{\partial x_m}\Big)\sin^{(k)}(2\pi f(x)), & \forall x \in S, \\ 0, & \text{otherwise,} \end{cases}$$

where c_β is a positive integer, $\alpha - \beta = (\alpha_1 - \beta_1, \ldots, \alpha_n - \beta_n)$ and $\beta \le \alpha$ iff $\beta_m \le \alpha_m$, $\forall m = 1, \ldots, n$. The term $(1 - \delta_{\beta_m,0})$ guarantees that only terms with $\beta_m > 0$ are considered.

The sum contains only terms of the form $t_1(x) = \partial^\beta(2\pi f(Ax))$ times $t_2(x) = \cos^{(k)}(2\pi f(Ax))$ for $x \in A^{-1}S$, respectively $t_1(x) = \partial^\beta(2\pi f(x))$ times $t_2(x) = \sin^{(k)}(2\pi f(x))$ for $x \in S$, where $1 \le |\beta| \le l$ and $1 \le k \le l$. By assumption, $t_1(x) = t_1(A^{-1}x) = t_1(A^{-2}x) = 0$, $\forall x \in S_\circ$ and the factor t_2 is bounded. Hence $\partial^\alpha h(x) = \partial^\alpha h(Ax) = \partial^\alpha h(A^2x) = 0$, $\forall x \in S_\circ$.

□

5.5.3 An example of a tight wavelet frame

EXAMPLE 5.5.1 (A radial tight wavelet frame):
Let us consider the problem of constructing a radial wavelet frame. Then the canonical choice for S is

$$S = \Big\{x \in \mathbb{R}^n : \frac{1}{4} \le |x| \le \frac{1}{2}\Big\}.$$

The problem to find a regular partition of unity of $C^l(\mathbb{R}^n)$-functions is now reduced to finding a function in $C^l(\frac{1}{4}, \frac{1}{2})$ such that $f(|\cdot|)$ satisfies the requirements of Theorem 5.5.7. We now look at examples of such functions f that give a regular partition of unity of $\mathbb{R}$ with respect to dyadic dilations. If $l \in \mathbb{N}_0$, we can easily find polynomials p_l satisfying the conditions $p_l(\frac{1}{4}) = 1/4$, $p_l(\frac{1}{2}) = 0$ and $p^{(k)}(m) = 0$, $\forall m = \frac{1}{4}, \frac{1}{2}$, $0 < k \le l$. Since a polynomial of degree one, namely $p_1(x) = 1/2 - x$ is required to satisfy $p_1(\frac{1}{4}) = 1/4$ and $p_1(\frac{1}{2}) = 0$, p_1 has a non-vanshing derivative, whence we require the polynomials to have degree $2l - 1$ to obtain $h \in C^l(\mathbb{R})$. Further examples for polynomials p_l resulting in functions $h \in C^l(\mathbb{R})$ are

$$\begin{aligned}
p_0(x) =& 1/2 - x \qquad (5.8)\\
p_1(x) =& -1 + 12x - 36x^2 + 32x^3\\
p_2(x) =& 8 - 120x + 720x^2 - 2080x^3 + 2880x^4 - 1536x^5\\
p_3(x) =& -52 + 1120x - 10080x^2 + 49280x^3 - 141120x^4 + 236544x^5 - 215040x^6 + 81920x^7\\
p_4(x) =& 368 - 10080x + 120960x^2 - 833280x^3 + 3628800x^4 - 10354176x^5 + 19353600x^6\\
&- 22855680x^7 + 15482880x^8 - 4587520x^9\\
p_5(x) =& -2656 + 88704x - 1330560x^2 + 11827200x^3 - 69189120x^4 + 279595008x^5 - 796207104x^6\\
&+ 1597685760x^7 - 2214051840x^8 + 2018508800x^9 - 1089994752x^{10} + 264241152x^{11}
\end{aligned}$$

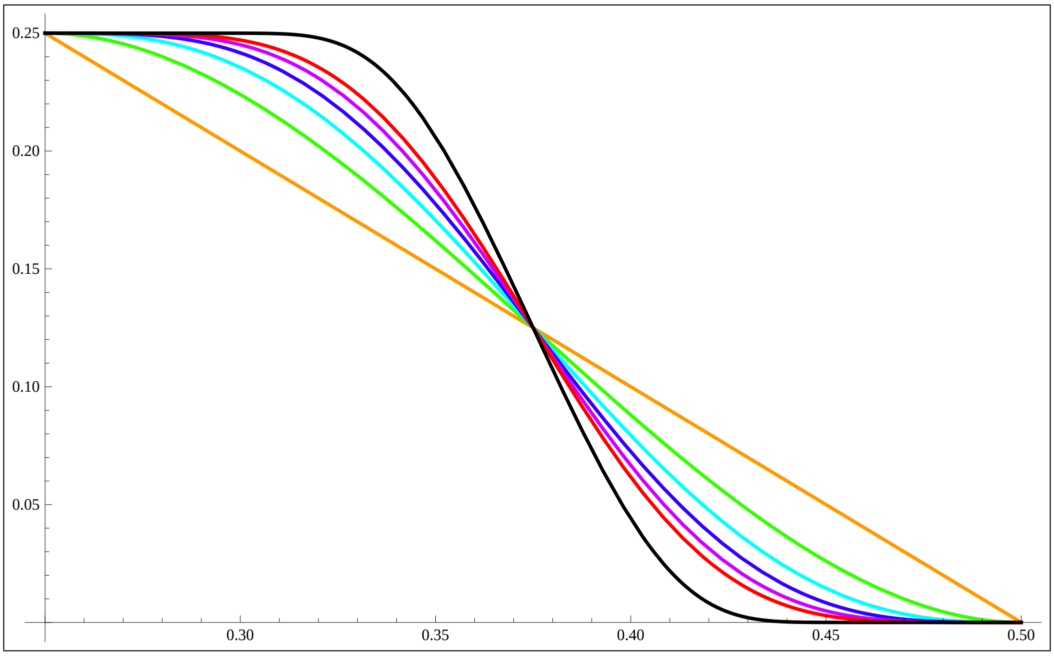


Figure 5.2: The functions p_0 (light brown), p_1 (green), p_2 (light blue), p_3 (dark blue), p_4 (violet), p_5 (red) and f_6 (black).

Now we give an example for a function $\lambda \in C^\infty(1,2)$ that yields a $h \in C^\infty(\mathbb{R})$.

Let $\lambda : \mathbb{R} \to \mathbb{R},\ x \mapsto \begin{cases} 0, & \forall x \leq 0 \\ e^{-x^{-2}}, & \forall x \in \mathbb{R}^+. \end{cases}$. Then $\lambda \in C^\infty(\mathbb{R})$ and $0 \leq \lambda(x) \leq 1\ \forall x \in \mathbb{R}^+$. If we let $\epsilon > 0$, then

$$f_\epsilon(x) = \frac{\lambda(\epsilon(\frac{1}{2} - x))}{4\big(\lambda(\epsilon(x - \frac{1}{4})) + \lambda(\epsilon(\frac{1}{2} - x))\big)}$$

is a suitable function.

5.5.4 An alternative construction for Riesz partitions of unity

In subsection 5.5.2 we provide a method for generating a regular partition of unity. Another way to generate a regular partition of unity is to transform an existing regular partition of unity on the real line with respect to translation $H = \{T_j h\}_{j \in \mathbb{Z}}$ to a regular partition of unity with respect to (dyadic) dilation. Every orthonormal system of translates of a single function ϕ in $L^2(\mathbb{R}^n)$ with compact support yields a suitable Riesz partition of unity – indeed $\sum_k |\widehat{\phi}(\xi + k)|^2 = 1$ holds almost everywhere on $\mathbb{R}$. (See [55] proposition 2.1.11 but note the different definition of the Fourier transform.) Of course for every multiresolution analysis there exists a scaling function $\phi \in L^2(\mathbb{R}^n)$ that yields an orthonormal system of translates. Thus, regular Riesz partitions of unity with respect to translation are easily available.

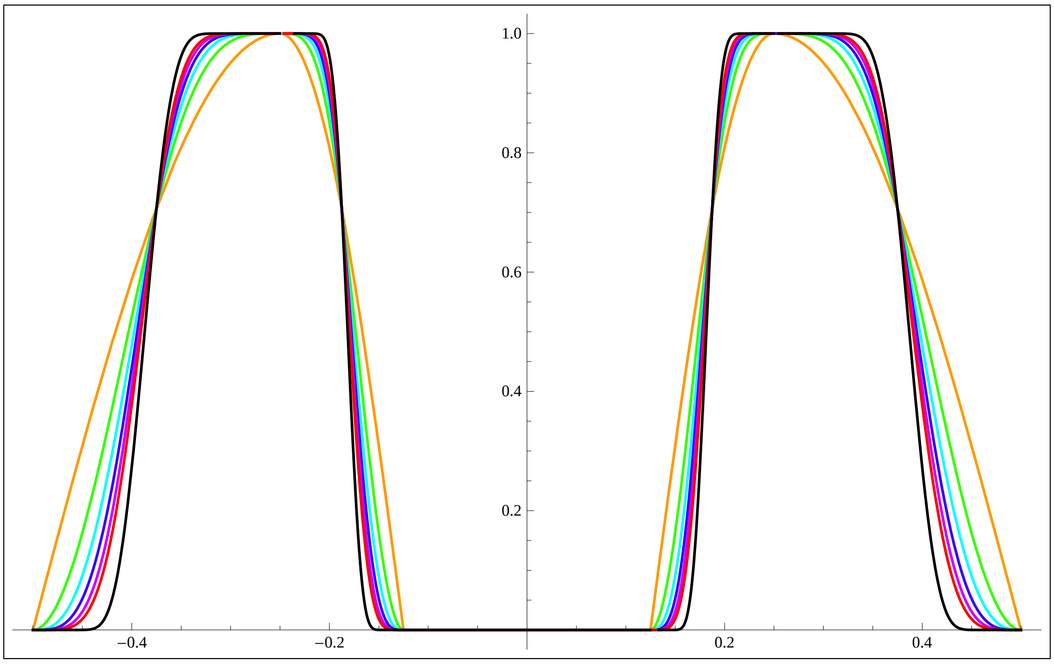


Figure 5.3: The function h corresponding to p_0 (light brown), p_1 (green), p_2 (light blue), p_3 (dark blue), p_4 (violet), p_5 (red) and f_6 (black).

Theorem 5.5.8
LET $h \in L^2(\mathbb{R})$ *be such that* $\operatorname{supp} h = [1, l]$ *and such that* $H := \{h_j\}_{j\in\mathbb{Z}} := \{T_j h\}_{j\in\mathbb{Z}}$, *is a regular partition of unity for* $\mathbb{R}$.
THEN $h' := \sum_{i=0}^{l} h(2^{-i}x + i)\chi_{[2^i,2^{i+1}]}(x)$ *defines a regular partition of unity of* $\mathbb{R}^+$ *with respect to the dyadic dilation by* $H' := \{h'(2^j\cdot)\}_{j\in\mathbb{Z}}$. *A regular partition of unity of* $\mathbb{R}^n \setminus \{0\}$ *is given by* $H' := \left\{h'(2^j|\cdot|)\right\}_{j\in\mathbb{Z}}$.

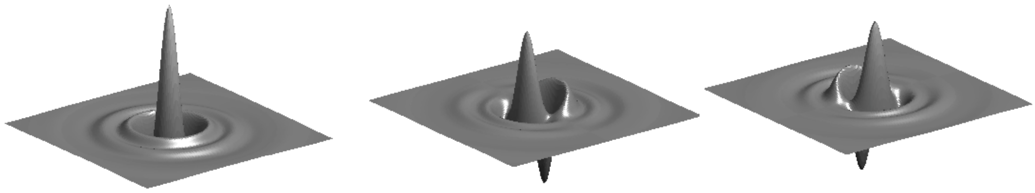
Proof. Let $x \in \mathbb{R}^+$. Then

$$\begin{aligned}\sum_{j\in\mathbb{Z}} \left(h'(2^j x)\right)^2 &= \sum_{j\in\mathbb{Z}} \sum_{i=0}^{l} \left(h(2^{j-i}x + i)\chi_{[2^i,2^{i+1}]}(2^j x)\right)^2 \\ &\overset{m=j-i}{=} \sum_{m\in\mathbb{Z}} \sum_{i=0}^{l} \left(h(2^m x + i)\chi_{[2^{-m},2^{-m+1}]}(x)\right)^2 \\ &\overset{*}{=} \sum_{i=0}^{l} \left(h(y+i)\right)^2 \\ &= 1,\end{aligned}$$

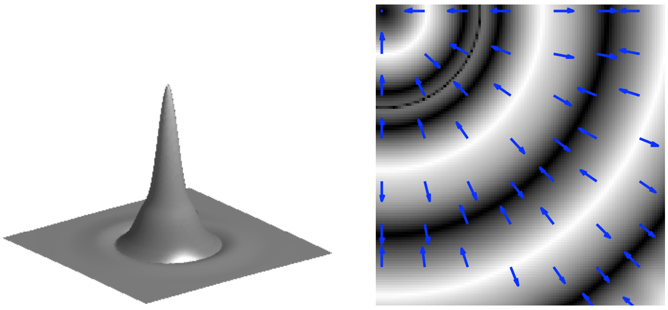
where $*$ holds with $y = 2^m x \in [1,2]$, with $m \in \mathbb{Z} : \chi_{[2^{-m},2^{-m+1}]}(x) = 1$. □

This theorem yields a regular partition of unity for $\mathbb{R}^n \setminus \{0\}$ by $h'(|x|)$.

The weakness of this method is that it does not conserve continuity or differentibility of the function $\widehat{\varphi}$. However, this is easily mended.



(a) Real wavelet (b) Riesz transformed wavelet



(c) Amplitude (d) Phase and phase direction

Figure 5.4: Wavelet for $L^2(\mathbb{R}^2)$ in the case $m = 2$. (a) Real part of the monogenic wavelet Ψ_m, i.e., the real wavelet Ψ. (b) Riesz transforms $R_1\Psi$, $R_2\Psi$ of the wavelet Ψ. (c) Amplitude $\|\Psi_m\|$. (d) Phase ϕ and phase direction (arrows) of the hypercomplex wavelet Ψ_m. The isotropy is clearly visible. The direction of the phase points from high ($\pi \leftrightarrow$ white) to low ($0 \leftrightarrow$ black) phase values.

Corollary 5.5.9
LET $h \in C^l(\mathbb{R})$ *be in compliance with the conditions of Theorem 5.5.8. Furthermore, let* $\tau : [1,2] \to [1,2]$ *such that* $\tau \in C^{(l)}(\mathbb{R})$, $\tau(1) = 1$, $\tau(2) = 2$ *and* $\tau^{(k)}(2) = (2^{1/2} D_{1/2}\tau)^{(k)}(2)\ \forall 1 \le k \le l$. THEN $h' := \sum_{i=0}^{l} h(2^{-i}\tau(x) + i)\chi_{[2^i,2^{i+1}]}(x)$ *is a regular partition of unity of* $\mathbb{R}^+$ *with respect to the dyadic dilation by* $H' := \{h'(2^j\cdot)\}_{j\in\mathbb{Z}}$. *Furthermore* $h' \in C^l(\mathbb{R}^+)$.

For an arbitrary dilation parameter a a Riesz partition of unity $G = \{g(a^j\cdot)\}_{j\in\mathbb{Z}}$ is given by choosing $g = h((a-1)(x+1)-1)$ and defining $g' = \sum_{i=0}^{l} h(a^{-i}x + (a-1)i)\chi_{[a^i,a^{i+1}]}(x)$.

5.5.5 Interpretation via refinable functions

Extension Principles

We now show that under certain conditions the wavelets that we constructed will satisfy the extension principle developed by Ron and Shen in [39]. It follows that these wavelets can be constructed by an extended concept of **MRA** (**Multi Resolution Analysis**). This yields the existence of a scaling function ϕ.

In order to state the extension principle, we need some notation.

Definition 5.5.10

LET *A be a dilation matrix.*
THEN

1. *a function $\phi \in L^2(\mathbb{R}^n)$ is called a* **refinable function**, *or* **scaling function**, *iff*

$$\exists H_0 \in L^2(\mathbb{T}) : \widehat{\phi}(A\xi) = H_0(\xi)\widehat{\phi}(\xi) \text{ a.e..}$$

   H_0 is called a **refinement mask**.

2. *the scaling function defines a space $V_0 := \overline{\operatorname{span}_{k\in\mathbb{Z}^n}\{T_k\phi\}}$. It follows that $V_j := A^j V_0 \subset V_{j-1}$ and $A^j\phi \in V_j$, $\forall j \in \mathbb{Z}$.*

3. *a set of functions $\Psi = \{\psi_l\}_{l=1}^L \subset L^2(\mathbb{R}^n)$ such that*

$$\forall l = 1,\ldots,L\, \exists H_l \in L^\infty(\mathbb{T}^n) : \widehat{\psi}_l(A\xi) = H_l(\xi)\widehat{\phi}(\xi) \text{ a.e.}$$

   is called a **wavelet system** *corresponding to the scaling function ϕ. The functions H_l are called* **wavelet masks**. *We denote $\mathfrak{H} := \begin{pmatrix} H_0 \\ H_1 \\ \vdots \\ H_L \end{pmatrix}$.*

4. *the* **fundamental function of multiresolution** *is defined as*

$$\Theta(\xi) := \sum_{r\in\mathbb{N}_0}\sum_{l=1}^{L} |H_l(A^r\xi)|^2 \prod_{j=0}^{r-1} |H_0(A^j\xi)|^2,\ \forall \xi \in \mathbb{R}^n \setminus \mathbb{Z}^n.$$

5. *the* κ-**function** *is defined by*

$$\kappa : \mathbb{R}^n \to \mathbb{Z},\ \xi \mapsto \inf\left\{j \in \mathbb{Z} :\ A^j\xi \in \mathbb{Z}^n\right\}.$$

The following is Theorem 6.5 of [39].

Theorem 5.5.11
LET *A be an integer valued dilation matrix. Let ϕ be a refinable function, Ψ a finite set of wavelets and H the corresponding refinement-wavelet mask. Assume that*

(i) ϕ satisfies the mild decay condition,

$$\sum_{\psi\in\Psi}\sum_{k\in\mathbb{N}} \Big\| \sum_{a\in A_k} |\widehat{\psi}(\,\cdot\, + a)|^2 \Big\|_{L^\infty([-1/2,1/2]^n)} < \infty,$$

where $A_k := \left\{a \in \mathbb{Z}^n : |a| > 2^k\right\}$; (This is satisfied if $\exists \rho > \frac{n}{2} :\ \widehat{\psi}(\xi) = O(\xi^{-\rho})$ as $\xi \to \infty$.)

(ii) $\widehat{\phi}(0) = \lim_{\xi\to 0}\widehat{\phi}(\xi) = 1$;

(iii) $\mathfrak{H}$ is essentially bounded.

THEN *Ψ generates a tight frame with bound C iff the following conditions hold*

1. *For a.e. $\xi \in \mathbb{R}^n$, $\lim_{j\to-\infty}\Theta(A^j\xi) = C$.*

2. *For a.e.* $\xi, \omega \in \mathbb{R}^n$, *if* $\kappa(\xi - \omega) = 1$, *then*

$$\Theta(A\xi) H_0(\xi) \overline{H_0(\omega)} + \sum_{l=1}^{L} H_l(\xi) \overline{H_l(\omega)} = 0$$

unless $\widehat{\phi}$ *vanishes identically on either* $\xi + \mathbb{Z}^n$ *or* $\omega + \mathbb{Z}^n$.

In particular, in the case $\Theta = 1$ *a.e.,* $\{A^{*j} T_k \psi_l\}_{j \in \mathbb{Z};\, k \in \mathbb{Z}^n,\, l=1,\dots,L}$ *is a tight frame if the vectors* $\mathfrak{H}$ *and* $T_\nu \mathfrak{H}$ *are perpendicular a.e., for every* $\nu \in (A^{-1}\mathbb{Z}^n / \mathbb{Z}^n) \setminus \{0\}$.

We now show that wavelet frames based on Theorem 5.5.7 satisfy the assumptions of the above theorem.

Corollary 5.5.12
LET *A be an integer valued dilation matrix and let* f *and* S *be as in Theorem 5.5.7. Set* $S_\cup := \cup_{j \in \mathbb{N}} A^{-j} S$.
THEN

$$g(\xi) := \begin{cases} 1, & \forall \xi \in A^{-1} S_\cup; \\ \sin(2\pi f(A\xi)), & \forall \xi \in A^{-1} S; \\ 0, & \text{otherwise,} \end{cases}$$

defines a refinable function ϕ *by* $\widehat{\phi} = g$.
Indeed a refinement mask is given by the 1*-periodic function defined on* $\mathbb{T}$ *by*

$$H_0(\xi) := \begin{cases} 1, & \forall \xi \in A^{-2} S_\cup; \\ \sin(2\pi f(A^2\xi)), & \forall \xi \in A^{-2} S; \\ 0, & \text{otherwise.} \end{cases}$$

The refinement mask satisfies $g(A\xi) = H_0(\xi) g(\xi)$.
The corresponding wavelet mask is given by

$$H_1(\xi) := \begin{cases} 0, & \forall \xi \in A^{-2} S_\cup; \\ \cos 2\pi f(A^2\xi), & \forall \xi \in A^{-2} S; \\ 1, & \forall \xi \in A^{-1} S; \\ 0, & \text{otherwise.} \end{cases}$$

The wavelet mask satisfies $h(A\xi) = H_1(\xi) g(\xi)$.

Proof. The fundamental function of multiresolution is in this case

$$\Theta(\xi) = \sum_{r \in \mathbb{N}_0} |H_1(A^r \xi)|^2 \prod_{j=0}^{r-1} |H_0(A^j \xi)|^2.$$

Now since $H_0(A^j \xi) = 1$, $\forall \xi \in \text{supp}\{H_0(A^{j-1}\xi), \forall j \in \mathbb{Z}\}$ the product reduces to

$$\prod_{j=0}^{r-1} |H_0(A^j \xi)|^2 = |H_0(A^{r-1}\xi)|^2.$$

Thus Θ is bounded, since

$$\begin{aligned}\Theta(\xi) &= \sum_{r\in\mathbb{N}_0} |H_1(A^r\xi)|^2 |H_0(A^{r-1}\xi)|^2 \\ &= \sum_{r\in\mathbb{N}_0} \begin{cases} 0, & \forall\xi\in A^{-r-1}S_\cup; \\ |\cos(2\pi f(A^{r+2}\xi))|^2, & \forall\xi\in A^{-r-1}S; \\ |\sin(2\pi f(A^{r+1}\xi))|^2, & \forall\xi\in A^{-r}S; \\ 0, & \text{otherwise.}\end{cases} \\ &= \begin{cases} 1, & \forall\xi\in S_\cup; \\ |\sin(2\pi f(A\xi))|^2, & \forall\xi\in S; \\ 0, & \text{otherwise.}\end{cases}\end{aligned}$$

As a consequence, condition 1. of Theorem 5.5.11 is met, since

$$\lim_{n\to-\infty}\Theta(A^n\xi) = 1,\ \forall\xi\in\mathbb{R}^n\setminus\{0\}.$$

To check condition 2., note that $\Theta(A\xi) = 1$, $\forall\xi\in\operatorname{supp}(H_0)$. Furthermore $\kappa(\omega-\xi) = 1$ is equivalent to $\xi-\omega\in A^{-1}\mathbb{Z}^n/\mathbb{Z}^n$. Hence we need to check if

$$H_0(\xi)\overline{H_0(\xi+q)} + H_1(\xi)\overline{H_1(\xi+q)} = 0,\ \forall\xi,\xi+q\in\operatorname{supp}(g),\ \forall q\in A^{-1}\mathbb{Z}^n\setminus\mathbb{Z}^n.$$

This condition is automatically fulfilled for the dyadic dilation $A = D_2$. This is because the dyadic dylation maps $\operatorname{supp} H_1 = D_2^{-1}S \subseteq [-\frac{1}{4},\frac{1}{4}]$ to $\operatorname{supp} H_1(D_2\cdot) = D_2^{-2}S \subseteq [-\frac{1}{8},\frac{1}{8}]$ and $|q|_\infty = \frac{2n+1}{2}$ holds for $q\in D_2^{-1}\mathbb{Z}^n/\mathbb{Z}^n$ and for some $n\in\mathbb{Z}\setminus\{1\}$. It follows that

$$\operatorname{supp} H_1(D_2^{-1}\cdot)\cap\operatorname{supp} T_q H_1(D_2^{-1}\cdot)$$

is a Lebesque zero set. Consequently

$$\operatorname{supp} H_0(D_2^{-1}\cdot)\cap\operatorname{supp} T_q H_0(D_2^{-1}\cdot) = \emptyset.$$

□

Isotropic multi resolution analysis

The method of construction for radial wavelet tight frames we derived in Theorem 5.5.7 can also be interpreted as a isotropic multiresolution analysis (IMRA). An introduction to IMRA can be found in [38]. In [38] a different method to generate radial wavelet frames is considered. However, in contrast to our method, this does not yield a wavelet frame with only one mother wavelet.

Definition 5.5.13 (IMRA)
An **isotropic multiresolution analysis** *(* **IMRA** *) of* $L^2(\mathbb{R}^n)$ *with respect to a rotated dilation* $\mathfrak{D}_a$ *whose matrix entries are integer valued is a sequence* $\{V_j\}_{j\in\mathbb{Z}}$ *of closed subspaces of* $L^2(\mathbb{R}^n)$ *satisfying the following conditions:*

(1) $\forall j\in\mathbb{Z},\ V_j\subset V_{j+1}$,

(2) $\mathfrak{D}_a^j V_0 = V_j$,

(3) $\bigcup_{j\in\mathbb{Z}} V_j$ *is dense in* $L^2(\mathbb{R}^n)$ *and* $\bigcap_{j\in\mathbb{Z}} V_j = \{0\}$,

(4) V_0 *is invariant under translations* $\{T_n\}_{n\in\mathbb{Z}^n}$,

(5) If P_0 *is the orthogonal projection onto* V_0, *then*

$$\rho^T P_0 = P_0 \rho,\ \forall \rho \in SO(n).$$

The next theorem gives conditions that are necessary for

$$\{\mathfrak{D}_a^j T_k \psi_l :\ k \in \mathbb{Z}^n,\ j \in \mathbb{N}_0,\ l = 1, \dots, L\} \cup \{T_k \phi :\ k \in \mathbb{Z}^n\}$$

to be a tight frame for $L^2(\mathbb{R}^n)$. It is cited from [38].

Theorem 5.5.14
LET $\phi \in L^2(\mathbb{R}^n)$ *be a refinable function with refinement mask* H_0 *such that* $\widehat{\phi}$ *is continuous at the origin and* $\lim_{|\xi|\to 0} \widehat{\phi}(\xi) = 1$. *Furthermore for* $l = 1, \dots, L$ *let* $H_l \in L^\infty(\mathbb{T})$ *be the wavelet masks for* $\psi_l \in L^2(\mathbb{R}^n)$.
THEN

$$\{\mathfrak{D}_a^j T_k \psi_l :\ k \in \mathbb{Z}^n,\ j \in \mathbb{N}_0,\ l = 1, \dots, K\} \cup \{T_k \phi :\ k \in \mathbb{Z}^n\}$$

is a tight frame for $L^2(\mathbb{R}^n)$ *with frame bound* $A = 1$ **if and only if** *for all* $q \in \mathfrak{D}_a^{-1}\mathbb{Z}^n / \mathbb{Z}^n$ *and for almost all* $\xi, \xi + q \in \mathbb{T}$ *such that*

$$\sum_{k\in\mathbb{Z}^n} |\widehat{\phi}(\xi + k)|^2 > 0 \wedge \sum_{k\in\mathbb{Z}^n} |\widehat{\phi}(\xi + q + k)|^2 > 0,$$

the equality

$$\sum_{l=0}^{L} H_l(\xi) H_l(\xi + q) = \delta_{q,0} \tag{5.9}$$

holds.

Proof. The proof can be found in [38] Theorem 4.7. □

The tight wavelet frames defined via Corollary 5.5.6 and Theorem 5.5.7 satisfiy all conditions but condition (5.9) by construction.
Since $\operatorname{supp}(\widehat{\phi}) = \mathfrak{D}_a^{-1} S$, we have that

$$\sum_{l=0}^{L} H_l^2(\xi) = H_0^2(\xi) + H_1^2(\xi) = 1,\ \forall \xi \in \operatorname{supp}(\phi) = \mathfrak{D}_a^{-1} S$$

Hence it is sufficient for the conditions of Theorem 5.5.14 to be met, that

$$\operatorname{supp} H_1(D_2^{-1}\cdot) \cap \operatorname{supp} T_q H_1(D_2^{-1}\cdot)$$

is a Lebesque zero set for any $q \in \mathfrak{D}_a^{-1}\mathbb{Z}^n / \mathbb{Z}^n$. It then follows that

$$\operatorname{supp} H_0(D_2^{-1}\cdot) \cap \operatorname{supp} T_q H_0(D_2^{-1}\cdot) = \emptyset.$$

This is satisfied in the case of dyadic dilation $\mathfrak{D}_a = D_2$ as seen above.

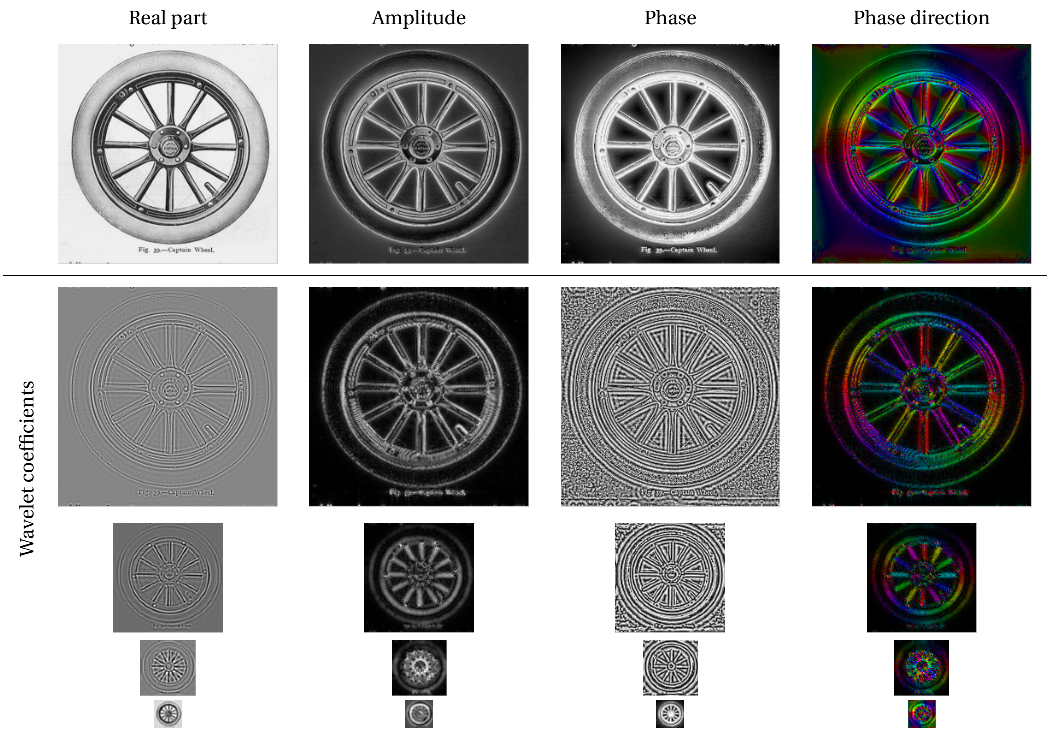


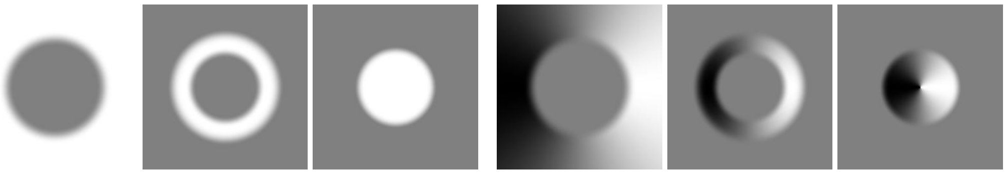
Figure 5.5: Example of a monogenic wavelet decomposition. First row: Original image amplitude, phase and phase direction of the monogenic signal. Second to fourth row: The first three scales of wavelet coefficients of the monogenic wavelet decomposition. Fifth row: Approximation coefficients (low pass component). The first column displays the real part of the wavelet coefficients, i.e., the e_0 component. The second and third column display the amplitude and the phase of the wavelet coefficients. The fourth column displays the color-coded phase direction: green points to the upper left and lower right corner, red encodes the horizontal direction, etc.. It is apparent that the phase direction displays the preferred directions on the corresponding scale. The descending sizes of the images represent the sub-sampling scheme of our filter-bank. The images show that the wavelet frame is extremely well adapted for round image features as well as for straight image features.

5.6 Implementation and application

The implementation of the monogenic wavelet has been the subject of the diploma thesis of Martin Storath [47] under supervision of the author. A more detailed description of the implementation can be found there.

5.6.1 Implementational aspects

The monogenic wavelet decomposition was implemented for 2-D and 3-D images as a plug-in for ImageJ [36]. For an example of the decomposition see Figure 5.5. The implementation uses the isotropic wavelets constructed in Example 5.5.1 for the orders $m = 0, \ldots, 5$ as given in (5.8) on page 95.



(a) Fourier filters h_2, h_1, h_0. (b) Fourier filters $r_1 h_2, r_1 h_1, r_1 h_0$.

Figure 5.6: Filter bank in the Fourier domain of the monogenic wavelets for the polynomial with $m = 3$ and $k = 2$ sub-channels. (a) Fourier transform h of the real wavelet Ψ from left to right: First sub-channel (high pass filter), second sub-channel and low pass filter. (b) Fourier transforms $\widehat{R_1\Psi}$, i.e., (a) multiplied by the Fourier multiplier of the Riesz transform $\frac{i\xi_1}{\|\xi\|}$. Here black $\leftrightarrow$ -1, gray $\leftrightarrow$ 0, and white $\leftrightarrow$ 1.

The filters for the Riesz transform and the wavelet frames are infinite impulse response (IIR) filters. Therefore, a classical implementation in space domain based on convolution would tend to be rather costly. Furthermore there would be accuracy problems due to the necessity of truncating filters.

Since both the Riesz transform and the wavelet frames do have a simple and explicit form in the Fourier domain, we implemented all algorithms in the frequency domain. The compact support of the wavelet frames in the Fourier domain allows a fast filter-bank implementation with perfect reconstruction property. The Shannon sampling theorem ensures loss-less down-sampling after each filter step. As we constructed tight wavelet frames, the synthesis filter bank consists of the same filters as the analysis filter bank.

The Fourier multiplier of the Riesz transform $\widehat{R_\alpha}(\xi) = \frac{i\xi_\alpha}{\|\xi\|}$ has a singularity at $\xi = 0$ (DC component). We deal with this singularity by subtracting the images mean value beforehand, which sets the DC component equal to zero.

For most applications, wavelet frames with dyadic dilations, as constructed above, are sufficient. Nevertheless, some applications deliver better results when a higher frequency resolution is used resulting in better separation of phase components. (See Example 5.1.1.) For that purpose, we constructed wavelet frames with $\sqrt[k]{2}$-expansive dilations, $k \in \mathbb{N}$, i.e., we use the dilation matrix $\sqrt[k]{2} I_n$. From a signal processing point of view, the high-pass Fourier filter corresponding to the dyadic case is divided into k sub-channels, which yields a $(k+1)$-channel filter bank ($k = 2$ in Fig. 5.6). We chose the $\sqrt[k]{2}$ dilations because we are then able to reuse our dyadic sub-sampling scheme after k filter steps.

The low-pass filter h_0, the band-pass filters h_j, $j = 1, \dots, k-1$, and the high-pass filter h_k are computed from the function h defined in Example 5.5.1 as follows:

$$\text{high-pass:} \qquad h_k(\xi) := \begin{cases} h(2^{k-1}\|\xi|^k), & \text{if } \|\xi\|^k 2^{k-1} < 1/4, \\ 1, & \text{if } \|\xi\|^k 2^{k-1} \geq 1/4, \end{cases}$$

$$\text{band-pass:} \qquad h_j(\xi) := h(2^{2k-j-1}\|\xi\|^k),$$

$$\text{low-pass:} \qquad h_0(\xi) := \begin{cases} 1, & \text{if } \|\xi\|^k 2^{2k-1} \leq 1/4, \\ h(2^{2k-1}\|\xi\|^k), & \text{if } \|\xi\|^k 2^{2k-1} > 1/4, \end{cases}$$

where $\xi \in [-1/2, 1/2]^n$.

Figure 5.7 desribes the first step of the filterbank algorithm for dimension $n = 2$, two subchannels ($k = 2$) and an $2^M \times 2^M$-image f, where $M \in \mathbb{N}$. A general step of the algorithm looks as

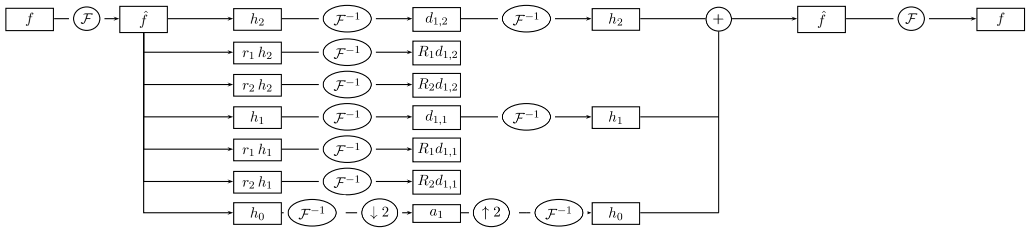


Figure 5.7: Filter bank for decomposition and reconstruction of signal a f by the monogenic wavelet frame. Here $k=2$ high pass filters are used, namely h_1 and h_2. h_0 is the approximation filter. (See Fig. 5.6 (a) for the filter masks.) $r_l h_j$ denotes the Fourier filter h_j multiplied by the Fourier multiplier of the Riesz transform $r_l = \frac{i\xi_l}{\|\xi\|}$ (see Fig. 5.6 (b) for case $l=1$). $d_{s,j}$ and $R_l d_{s,j}$ are the real and hypercomplex detail coefficients, respectively. The first index s denotes the scale of the coefficients.

follows:

EXAMPLE 5.6.1 (A single step of the algotithm):
Application of the wavelet filters h_l of octave $l = 1,2$ in the Fourier domain yields the wavelet coefficients of the corresponding octave $d_{s,l}$. Application of the Riesz transformed filters yields the Riesz transformed wavelet octave filters $R_1 d_{s,l}, R_2 d_{s,l}$. Application of the low pass filter h_0 yields the approximation component a_s which is not Riesz transformed since it contains the DC-component. After downsampling the approximation component is the initial point for the next step of the filter bank algorithm. The amplitude of the wavelet component $A_{s,l}$ is computed as

$$A_{s,l} = \sqrt{d_{s,l}^2 + R_1 d_{s,l}^2 + R_2 d_{s,l}^2}.$$

The phase $\phi_{s,l}$ is calculated as

$$\phi_{s,l} = \arg\left(d_{s,l} + i\sqrt{R_1 d_{s,l}^2 + R_2 d_{s,l}^2}\right).$$

It follows that

$$d_{s,l} = A_{s,l}\cos(\phi_{s,l}).$$

The phase direction $D_{s,l}$ is computed via

$$D_{s,l} := \begin{cases} \arg\left(\frac{R_1 d_{s,l}}{\sqrt{R_1 d_{s,l}^2 + R_2 d_{s,l}^2}} + i\frac{R_2 d_{s,l}}{\sqrt{R_1 d_{s,l}^2 + R_2 d_{s,l}^2}}\right), & \text{if } R_2 d_{s,l} \geq 0; \\ \arg\left(\frac{R_1 d_{s,l}}{\sqrt{R_1 d_{s,l}^2 + R_2 d_{s,l}^2}} + i\frac{R_2 d_{s,l}}{\sqrt{R_1 d_{s,l}^2 + R_2 d_{s,l}^2}}\right) + \pi, & \text{else.} \end{cases}$$

The runtime of the monogenic wavelet decomposition is determined by that of the fast Fourier transform, namely $O(N \log N)$. The memory consumption amounts to a factor of less than $2(n+1)$ times the image size, where $n=2$ or $n=3$ is the image dimension. In the case of $\sqrt[k]{2}$-expansive dilations, the runtime and the redundancy increase by a factor of k as compared to the dyadic case. E.g., for the application "Equalization of Brightness" (see section 5.6.2), $k=5$ turned out to be a sensible trade-off between frequency resolution and runtime.

We computed images of different sizes with the ImageJ implementation of our fast algorithm. The runtime and accuracy results of the experiments are given in Table 5.1. The examples were

Table 5.1: Computation time and accuracy of the Riesz wavelet filter bank for various sample sizes. The time for the creation of the filter-bank is denoted as pre-calculation time. The reconstruction error lies in the range of machine accuracy. The Examples were computed on an Intel Core 2 Duo with 2.33 GHz.

Sample Size	Precalculation Time in sec	Analysis in sec	Synthesis in sec	l^∞-Reconstruction Error	Equalization of brightness in sec
128 × 128	0.02	0.03	0.05	$4.2\cdot10^{-16}$	0.19
256 × 256	0.10	0.10	0.14	$5.1\cdot10^{-16}$	1.63
512 × 512	0.40	0.45	0.56	$6.7\cdot10^{-16}$	6.66
1024 × 1024	2.00	1.96	2.36	$9.1\cdot10^{-16}$	28.67
128×128×128	3.45	2.66	4.20	$3.9\cdot10^{-16}$	71.43

calculated using dyadic dilations. In section 5.6.2 below, we used $\sqrt[5]{2}$-expansive dilations. Note that the filters have to be computed only *once* because they are used both for decomposition and reconstruction.

5.6.2 Applications

Equalization of Brightness

Looking at the monogenic wavelet decomposition we, observe that the amplitude contains information about the brightness of the image, whereas the phase contains the structure of the image. Note that this observation is true only for the amplitude and phase of the wavelet coefficients of the image, and not for the amplitude and phase decomposition of the image itself. (Compare Fig. 5.11.)

Hence, reconstruction by phase only delivers an image with balanced brightness (Fig. 5.8). This reconstruction by phase only is achieved by setting the amplitude of the wavelet coefficients to be constant. Details are conserved or even enhanced by equalization of brightness. As mentioned above the results become better when using $\sqrt[k]{2}$-expansive dilations instead of dyadic ones (we chose $k = 5$). The algorithm also sets very low amplitudes to one resulting in an amplification of noise. Therefore, we added a simple noise suppression based on the magnitude of the amplitude.

Equalization of brightness can also be applied to color images. See [47] for the details.

The multi-scale phase concept is closely connected to the human visual perception. This can be seen in the example of monogenic bandpass filters [16] applied to Adelsons checkerboard illusion [1]. In Figure 5.10, we reproduce this effect using our monogenic wavelet frame. Note that in contrast to [16] no structural information is lost since we use a wavelet frame instead of a set of bandpass filters. Hence the resulting images are far closer to the original images. (Compare Figure 5.10 with figure 1 in [16].)

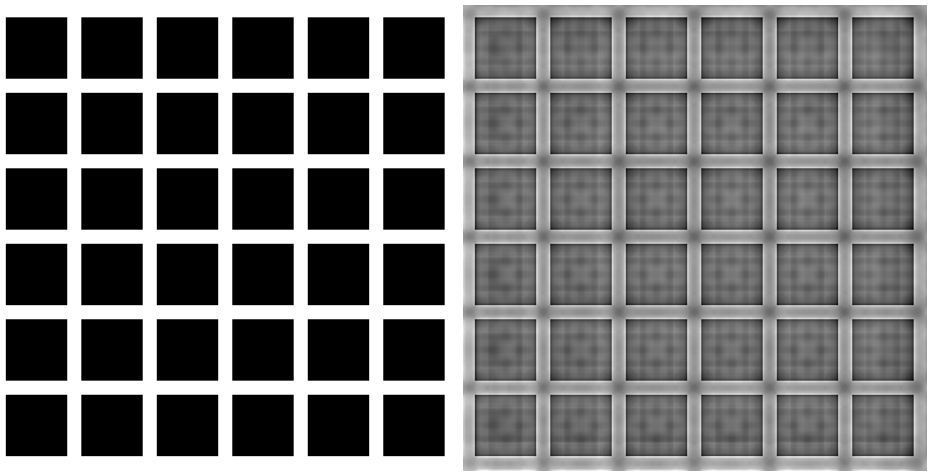
Another hint that the human visual perception is related to a multi-scale phase concept is provided by the optical illusion in Figure 5.9. The image consists of black squares separated by white stripes, but the human vision injects small gray disks into the intersection of the horizontal and vertical stripes. This phenomenon is reproduced in the phase image derived by equalization of brightness.

(a) Corrupted Image (b) Equalization of Brightness

Figure 5.8: (a) Artificially corrupted image. Gray levels are multiplied by a ramp function and by 0.05 in the center. (b) Reconstruction from phase only. The brightness is now balanced, in particular the difference of brightness between the dark center and its outside disappears. Even fine image structures are recovered. Decomposition was computed using $k = 5$ sub-channels.

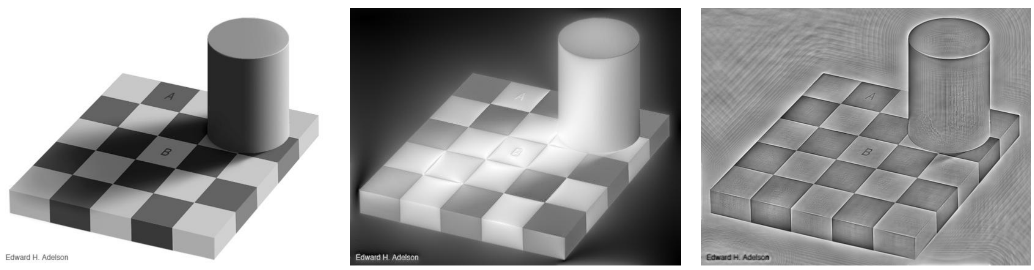
Parameter Free Descreening

An application using the amplitude of the monogenic wavelet coefficients is parameter free descreening. Screening effects in images are usually removed by applying an adequate low pass filter. The filter size has to be given as a parameter or estimated in some way. We use the fact that due to the regularity of the screening effect the maximal amplitude lies in the subband corresponding to the screening effect to automatically find the adequate low pass filter that removes the screening [24, 50].



(a) Image with optical illusion (b) Equalization of brightness

Figure 5.9: The optical illusion of injecting gray disks is reproduced by the phase image (b) derived by Equalization of Brightness.



(a) Original (b) Phase (c) Reconstruction from Phase

Figure 5.10: (a) Adelsons checkerboard illusion [1]: Squares A and B seem to have different brightness, but in fact they have the same gray levels. (b) In the phase image, A and B *have* different gray levels. The contour of the shadow is still visible. (c) After the reconstruction from phase of wavelet scales, A and B have again different gray levels and the contour of the shadow disappeared. Regarding the structure of the image, the human vision is in some way related to image (c): The absolute gray values in image (a) do *not* yield a checkerboard structure, as A and B have the same gray values. But we recognize (a) in the way of (c): Our visual system compensates the shadow that is thrown by the cylinder. (Compare also [16].)

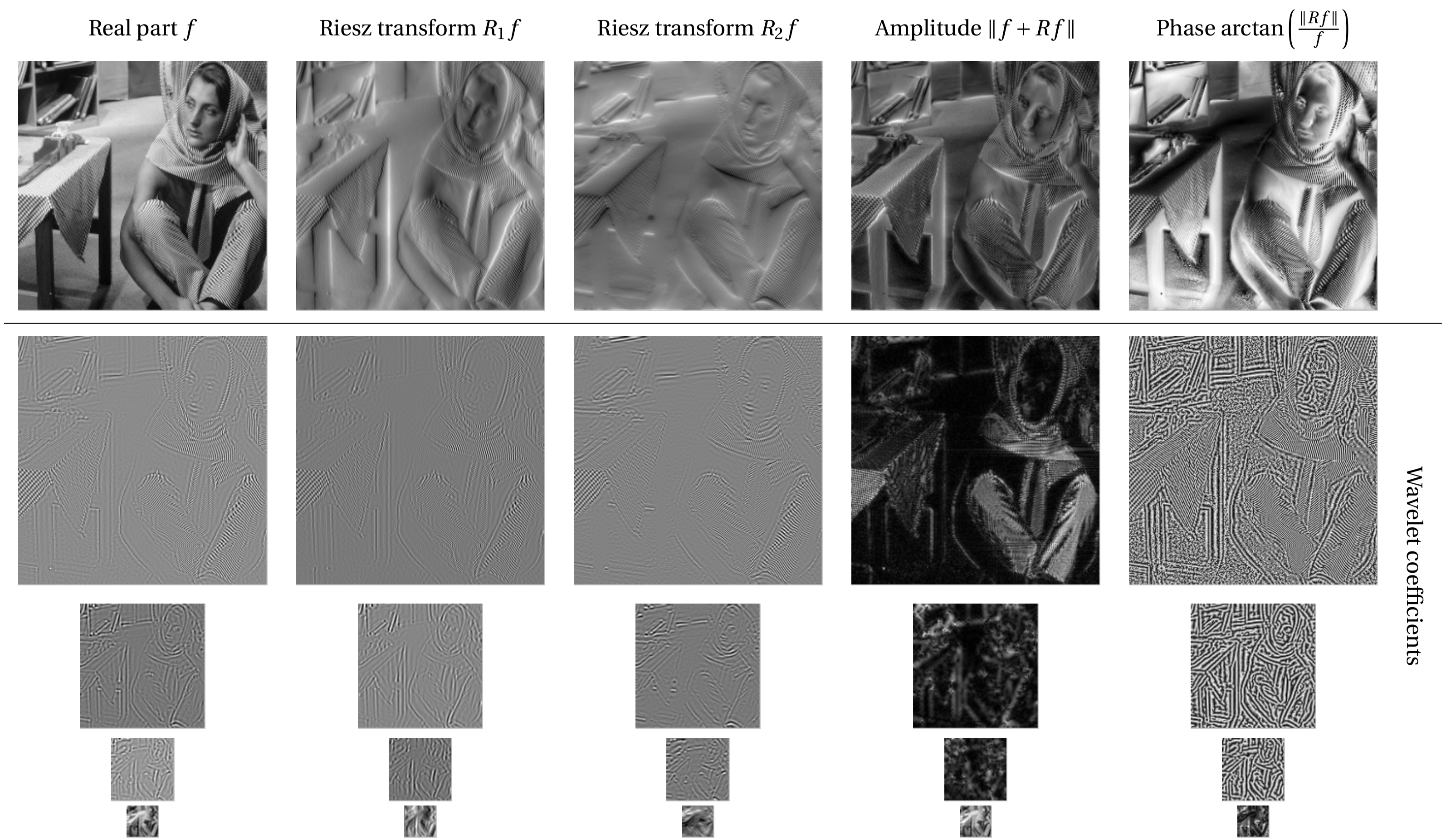


Figure 5.11: First row: Original image and Riesz transforms in both directions. Second to fourth row: The first three scales of the wavelet coefficients of the monogenic wavelet decomposition (band pass component). Fifth row: Approximation coefficients (low pass component). The first column displays the real part of the wavelet coefficients, i.e., the e_0 component, the second and third column the imaginary parts, i.e., the e_1 and e_2 components. We observe that the Riesz transform acts as a directional filter. (See, e.g., the table leg in the image.) The fourth and fifth column are the amplitude and the phase of the wavelet coefficients. The descending sizes of the images represent the sub-sampling scheme of our filter-bank. Notice that the amplitude of the wavelet coefficients of the image is indeed only large where there are image features of the corresponding scale. In contrast the amplitude of the original signal has no such obvious interpretation.

Chapter 6

Uncertainty relations for the Riesz transform

Uncertainty relations quantify the joint localisation in time, or space domain, and Fourier domain. For that reason they are an important tool in the design of wavelets. In chapter 5 we implement Hilbert and Riesz transform via wavelets. Therefore it will be the topic of this chapter to compute the effect of the Hilbert and Riesz transforms on the Weyl-Heisenberg uncertainty relation.

In Theorem 6.2.3 we proof the new result that under mild assumptions the Weyl-Heisenberg uncertainty relation is invariant under the Hilbert transform.

Since the Riesz transform R is unitary (See Theorem 4.1.4.) whereas the partial Riesz transforms R_α are not it seems reasonable to consider a single uncertainty relation for the Riesz transform rather than a set of uncertainty relations for the partial Riesz transforms. This requires a new kind of uncertainty relation for vector valued functions.

We pursue this goal along two different paths: A classical approach Theorem 6.3.2 based on the Cauchy-Schwartz inequality and an uncertainty relation on Clifford-Hilbert modules Theorem 6.3.3. Both paths yield novel uncertainty relations.

In contrast to the set of classical Weyl-Heisenberg uncertainty relations in higher dimensions stated in Corollary 6.3.1 the Weyl-Heisenberg uncertainty relation we derive in Theorem 6.3.10 is invariant under rotations. This property of invariance with respect to rotation is very desirable for image processing.

In Theorem 6.3.12 we use our results to derive a novel Weyl-Heisenberg uncertainty relation for the Riesz transform and to compute the effect of the Riesz transform on the Weyl-Heisenberg uncertainty relation stated in Theorem 6.3.10.

Finally we state a new affine uncertainty relation Theorem 6.3.13 as a further example for an application of Theorem 6.3.3.

Part of the results of this chapter have been published in the peer reviewed article [22].

6.1 The Heisenberg algebra and the Weyl Heisenberg group

Definition 6.1.1 (Heisenberg algebra and Weyl-Heisenberg group)
LET $a, b, c, d \in \mathbb{R}^n$ *and* $s, t \in \mathbb{R}$.
THEN $(a, b, t), (c, d, s) \in \mathbb{R}^{2n+1}$ *and we define a* **Lie bracket** *on* $\mathbb{R}^{2n+1}$ *by*

$$[(a, b, t), (c, d, s)] = (0, 0, ad - bc).$$

The **Lie algebra** $h_n := (\mathbb{R}^{2n+1}, [\,\cdot\,,\,\cdot\,])$ *is called the* **Heisenberg algebra**. *(Lie algebras and Lie groups are defined in the Appendix, subsection A.1.1.)*
If $\{A_1, \dots, A_n, B_1, \dots, B_n, T\}$ *is the standard basis of* $\mathbb{R}^{2n+1}$ *then*

$$[A_j, A_k] = [B_j, B_k] = [A_j, T] = [B_j, T] = 0, \text{ and } [A_j, B_k] = \delta_{j,k} T.$$

A **representation of the Heisenberg algebra** *on* $\mathscr{S}(\mathbb{R}^n)$ *is given by*

$$d\mathscr{W}_{a,b,t} f(x) := (a \cdot \nabla + 2\pi i b \cdot x + 2\pi i t \,\mathrm{Id}) f(x), \; \forall f \in \mathscr{S}(\mathbb{R}^n), \; x \in \mathbb{R}^n.$$

That is the mapping from h_n *to the set of skew-Hermitian operators* $\{d\mathscr{W}_{a,b,t}\}_{a,b \in \mathbb{R}^n,\, t \in \mathbb{R}}$ *is a Lie algebra homomorphism.*
The **Weyl-Heisenberg group** H_n *is* $\mathbb{R}^{2n+1}$ *with group law*

$$(a, b, t)(c, d, s) = (a + c, b + d, t + s + 1/2(ad - bc)).$$

The Weyl-Heisenberg group is the **Lie group** *corresponding to the Heisenberg Lie algebra.*
The **Schrödinger representation** *of* H_n *is given by the set of unitary operators* $\{\mathscr{W}_{a,b,t}\}_{a,b \in \mathbb{R}^n,\, t \in \mathbb{R}}$, *where*

$$\mathscr{W}_{a,b,t} : L^2(\mathbb{R}) \to L^2(\mathbb{R}), \; f \mapsto e^{2\pi i (t+ab)} M_b T_a f.$$

The **infinitesimal operators** *of a group representation are the derivatives at the identity of the operator defining the representation with respect to the basis elements of the group represented. The infinitesimal operators of a Lie group are a basis for a representation of the corresponding Lie algebra.*

Proposition 6.1.2 (Infinitesimal operators of the Weyl-Heisenberg group)
The infinitesimal operators of the Weyl-Heisenberg group are

$$\begin{aligned}
t_\alpha f(x) &:= \frac{d}{da_\alpha} e^{2\pi i (t+ab)} M_b T_a f(x)\big|_{a=b=0,\, t=0} \\
&= 2\pi i b_\alpha e^{2\pi i (t+ab)} M_b T_a f(x) - e^{2\pi i (t+ab)} M_b T_a \frac{\partial}{\partial x_\alpha} f(x)\big|_{a=b=0,\, t=0} \\
&= -\frac{\partial}{\partial x_\alpha} f(x), \\
s_\alpha f(x) &:= \frac{d}{db_\alpha} e^{2\pi i (t+ab)} M_b T_a f(x)\big|_{a=b=0,\, t=0} \\
&= 2\pi i a_\alpha e^{2\pi i (t+ab)} M_b T_a f(x) + 2\pi i x_\alpha e^{2\pi i (t+ab)} M_b T_a f(x)\big|_{a=b=0,\, t=0} \\
&= 2\pi i x_\alpha f(x), \\
&\frac{d}{dt} e^{2\pi i (t+ab)} M_b T_a f\big|_{a=b=0,\, t=0} = 2\pi i f.
\end{aligned}$$

As we have noted above these operators are a basis for the representation of the Heisenberg algebra on $\mathscr{S}(\mathbb{R}^n)$.

6.2 Uncertainty relations in one dimension

Uncertainty principles in one dimension are based on a uncertainty relation for self-adjoint operators stated in this form for example in [12].

Theorem 6.2.1
LET *H be a Hilbert space and let S, T be densely defined self-adjoint operators*

$$S : H \supseteq \mathrm{Dom}(S) \to H,$$
$$T : H \supseteq \mathrm{Dom}(T) \to H.$$

Furthermore, let the commutator $[S,T]$ be given by

$$[S,T] := ST - TS.$$

THEN *for any function $f \in H$ and any pair of real numbers $s,t \in \mathbb{R}$ the following inequality holds:*

$$\left|\langle [S,T]f, f\rangle\right|^2 \leq 4\|(S-s)f\|^2\|(T-t)f\|^2. \tag{6.1}$$

Equality in (6.1) holds iff

$$\exists \lambda \in \mathbb{R} : (S-s)f = i\lambda(T-t)f.$$

If $\|f\| = 1$, then equality in (6.1) implies that

$$s = \langle Sf, f\rangle, \text{ and } t = \langle Tf, f\rangle.$$

Furthermore (6.1) can be rewritten as

$$\left|\langle [S,T]f, f\rangle\right|^2 \leq 4\langle (T-t)^2 f, f\rangle\langle (S-s)^2 f, f\rangle$$

$\langle (T-t)^2 f, f\rangle$ is called the variance of T.

Proof. A proof can be found in [12]. □

The operators $S_\alpha = is_\alpha$ and $T_\beta = it_\beta$ in Proposition 6.1.2 are self-adjoint and have a non-vanishing commutator iff $\alpha = \beta$. Hence we can derive an uncertainty relation of them via Theorem 6.2.1 - the Heisenberg uncertainty in one dimension. (See for example [20] which contains a beautiful introduction to uncertainty relations.)

Theorem 6.2.2 (Heisenberg uncertainty relation)
LET *$f \in L^2(\mathbb{R})$, and let $l, r \in \mathbb{R}$.*
THEN

$$16\pi^2 \int_{\mathbb{R}} (x+r)^2 |f(x)|^2 dx \int_{\mathbb{R}} (\xi + l)^2 |\widehat{f}(\xi)|^2 d\xi \geq \|f\|^4. \tag{6.2}$$

Equality holds if $f(x) = e^{-\pi\frac{x^2}{\lambda}}$, for some $\lambda \in \mathbb{R}$.

Proof. Let $Sf(x) := -2\pi x f(x)$ and $Tf(x) := -if'(x)$. S and T are densely defined self-adjoint operators on $L^2(\mathbb{R})$. Furthermore let $s = 2\pi r$ and $t = 2\pi l$. We can apply (6.2.1) which states that

$$4\|(S-s)^2 f\|^2 \|(T-t)^2 f, f\|^2 \geq \left|\left\langle [S-s, T-t]f, f\right\rangle\right|^2.$$

Then

$$\begin{aligned}[S-s,T-t] &= (S-s)(T-t)-(T-t)(S-s)\\ &= ST-TS-sT+Ts-St+tS+st-ts\\ &= [S,T].\end{aligned}$$

Hence, the right hand side of the inequality (6.2.1) yields

$$\begin{aligned}\left|\langle[(S-s),(T-t)]f,f\rangle\right|^2 &= \left|\int_{\mathbb{R}} 2\pi i x f'(x) - 2\pi i \frac{\partial}{\partial x}(xf(x))\overline{f(x)}dx\right|^2\\ &= \left|\int_{\mathbb{R}} 2\pi i\big(xf'(x)-f(x)-xf'(x)\big)\overline{f(x)}dx\right|^2\\ &= \left|2\pi i\right|^2\|f\|^4.\end{aligned}$$

The first term of the left hand side of (6.2.1) is simply

$$\langle(S-s)^2 f,f\rangle = \int(-2\pi x-s)^2 f(x)\overline{f(x)}dx = \int(-2\pi x-s)^2|f(x)|^2dx.$$

The second term gives

$$\begin{aligned}\langle(T-t)^2 f,f\rangle = \langle(T-t)f,(T-t)f\rangle = \|(T-t)f\|^2 &= \left\|\mathscr{F}\big((T-t)f\big)\right\|^2\\ &= \int_{\mathbb{R}}(-2\pi x-t)\widehat{f}(x)\overline{(-2\pi x-t)\widehat{f}(x)}dx\\ &= \int_{\mathbb{R}}(-2\pi x-t)^2|\widehat{f}(x)|^2dx.\end{aligned}$$

Equivalence holds if $\exists\lambda\in\mathbb{R}:(S-s)f=i\lambda(T-t)f$. This is clearly true if $f(x)=e^{-\pi\frac{x^2}{\lambda}}$. □

6.2.1 The Heisenberg uncertainty and the Hilbert transform

The following new theorem states that the Heisenberg uncertainty relation introduced in Theorem 6.2.2 is invariant under the Hilbert transform if the first moment vanishes.

Theorem 6.2.3 (Heisenberg uncertainty relation and the Hilbert transform)
LET $f\in L^2(\mathbb{R},\mathbb{R})$ *such that* $Sf,Tf\in L^2(\mathbb{R})$ *and additionally* $\int_{\mathbb{R}} f(x)dx=0$.
THEN

$$\int_{\mathbb{R}} x^2|f(x)|^2dx\int_{\mathbb{R}}\xi^2|\widehat{f}(\xi)|^2d\xi = \int_{\mathbb{R}} x^2|\mathscr{H}f(x)|^2dx\int_{\mathbb{R}}\xi^2|\widehat{\mathscr{H}f}(\xi)|^2d\xi.$$

Proof. We first show that $S\mathscr{H}f\in L^2(\mathbb{R})$. Indeed we show that $T\widehat{\mathscr{H}f}\in L^2(\mathbb{R})$. First note that if $Sf\in L^2(\mathbb{R})$ then $\widehat{f}$ is an element of the Sobolev space

$$W^{1,2}(\mathbb{R}) := \{g\in L^2(\mathbb{R}): g'\in L^2(\mathbb{R})\},$$

with norm $\|.\|_{W^{1,2}}$ given by $\|g\|_{W^{1,2}} = \|g\|_{L^2(\mathbb{R})}+\|g'\|_{L^2(\mathbb{R})}$, $\forall g\in W^{1,2}(\mathbb{R})$. By the Sobolev embedding theorem (See [26] theorem 4.5.11) $\widehat{f}\in C(\mathbb{R})$ and consequently $\widehat{f}(0)=0$.
Let $\{f_k\}_{k\in\mathbb{N}}\subset\mathbb{C}^\infty(\mathbb{R})$ such that $f_k(0)=0$ and

$$f_k \overset{k\to\infty}{\longrightarrow} \widehat{f} \text{ in } W^{1,2}(\mathbb{R}).$$

In the distributional sense $\operatorname{sgn}' = 2\delta_0$ and hence

$$
\begin{aligned}
(f_k \operatorname{sgn})'(\phi) &= f_k' \operatorname{sgn}(\phi) + 2 f_k \delta_0(\phi) \\
&= \int_{\mathbb{R}} f_k'(x) \operatorname{sgn}(x) \phi(x)\, dx + \underbrace{2 f_k(0)\phi(0)}_{=0}, \ \forall \phi \in \mathcal{D}(\mathbb{R}).
\end{aligned}
$$

It follows from the Cauchy-Schwarz inequality that

$$
\begin{aligned}
&\left| f_k' \operatorname{sgn}(\phi) - \widehat{f}' \operatorname{sgn}(\phi) \right| \\
&\quad = \left| \int_{\mathbb{R}} \left(f_k'(x) \operatorname{sgn}(x) - \widehat{f}'(x) \operatorname{sgn}(x) \right) \phi(x)\, dx \right| \\
&\quad \le \sqrt{\int_{\mathbb{R}} \left| \left(f_k'(x) - \widehat{f}'(x) \right) \operatorname{sgn}(x) \right|^2 dx} \sqrt{\int_{\mathbb{R}^n} |\phi(x)|^2 dx} \\
&\quad \xrightarrow{k\to\infty} 0, \ \forall \phi \in \mathcal{D}(\mathbb{R})
\end{aligned}
$$

This shows that

$$
\widehat{\mathcal{H} f}' = i \operatorname{sgn} \widehat{f}' \in L^2(\mathbb{R}). \tag{6.3}
$$

We can now compute

$$
\begin{aligned}
&\int_{\mathbb{R}} x^2 |\mathcal{H} f(x)|^2 dx \int_{\mathbb{R}} \xi^2 |\widehat{\mathcal{H} f}(\xi)|^2 d\xi \\
&= \int_{\mathbb{R}} \frac{1}{4\pi^2} \left| \frac{\partial}{\partial \xi} \widehat{\mathcal{H} f(\xi)} \right|^2 d\xi \int_{\mathbb{R}} \xi^2 |i \operatorname{sgn}(\xi) \widehat{f}(\xi)|^2 d\xi \\
&\overset{(6.3)}{=} \frac{1}{4\pi^2} \left(\int_{\mathbb{R}^+} \left| i \operatorname{sgn}(\xi) \widehat{f}'(\xi) \right|^2 d\xi + \int_{\mathbb{R}^-} \left| i \operatorname{sgn}(\xi) \widehat{f}'(\xi) \right|^2 d\xi \right) \int_{\mathbb{R}} \xi^2 |\widehat{f}(\xi)|^2 d\xi \\
&= \frac{1}{4\pi^2} \int_{\mathbb{R}} \left| \widehat{f}'(\xi) \right|^2 d\xi \int_{\mathbb{R}} \xi^2 |\widehat{f}(\xi)|^2 d\xi \\
&= \int_{\mathbb{R}} x^2 |f(x)|^2 dx \int_{\mathbb{R}} \xi^2 |\widehat{f}(\xi)|^2 d\xi.
\end{aligned}
$$

□

6.3 Uncertainty in higher dimensions

We have just seen that the Hilbert transform does not change the localization of a function, i.e., the Heisenberg uncertainty relation is invariant under the Hilbert transform. If we want to know if this also applies to the Riesz transform, we have to compute the Heisenberg uncertainty of the Riesz transform of a function and compare it to the Heisenberg uncertainty of the original function. Therefore we derive an uncertainty relation for vector valued functions like the Riesz transform of a function.

The uncertainty principle Theorem 6.2.2 is usually generalized to higher dimensions by applying Theorem 6.2.1 to the partial infinitesimal operators calculated in Proposition 6.1.2. (See for example [17] Corollary 2.6.)

Corollary 6.3.1
LET $f \in L^2(\mathbb{R}^n)$, *and* $l = (l_1, \ldots, l_n), r = (r_1, \ldots, r_n) \in \mathbb{R}^n$.
THEN

$$
16\pi^2 \int_{\mathbb{R}^n} (x_\alpha + r_\alpha)^2 |f(x)|^2 dx \int_{\mathbb{R}^n} (\xi_\alpha + l_\alpha)^2 |\widehat{f}(\xi)|^2 d\xi \ge \|f\|^4, \ \forall \alpha = 1, \ldots, n. \tag{6.4}
$$

Corollary 6.3.1 states a set of n uncertainty relations. We suggest two ways to get a single uncertainty relation.

In Theorem 6.3.2, we first derive a uncertainty relation from Corollary 6.3.1.

In Theorem 6.3.3 we consider the infinitesimal operators derived from the n-dimensional Weyl-Heisenberg group by application of Dirac operators of the form $\sum_{\alpha=1}^{n} e_\alpha \frac{\partial}{\partial x_\alpha}$ to obtain a single uncertainty relation, rather than partial derivatives of the form $i\frac{\partial}{\partial x_\alpha}$ to get n uncertainty relations. This results in an uncertainty relation for operators in Clifford-Hilbert modules which is more general but also a lot weaker than that in Theorem 6.3.2.

6.3.1 A single uncertainty relation for sets of self adjoint operators

Theorem 6.3.2
LET S_α, T_α, $\alpha = 1,\dots,n$, *be densely defined self-adjoint operators*

$$S_\alpha : H \supseteq \mathrm{Dom}(S_\alpha) \to H,$$
$$T_\alpha : H \supseteq \mathrm{Dom}(T_\alpha) \to H,$$

such that $\cap_\alpha \mathrm{Dom}(S_\alpha)$ *and* $\cap_\alpha \mathrm{Dom}(T_\alpha)$ *are dense in* H.

THEN *for any set of functions* $f_1,\dots,f_m \in \cap_{\alpha=1}^{n} \mathrm{Dom}([S_\alpha, T_\alpha])$ *and any pair of vectors* $(s_1,\dots,s_n),(t_1,\dots,t_n) \in \mathbb{R}^n$ *the following inequality holds:*

$$\Big(\sum_{\alpha=1}^{n}\sum_{k=1}^{m}\big|\langle[S_\alpha,T_\alpha]f_k,f_k\rangle\big|\Big)^2 \le 4\sum_{\alpha=1}^{n}\sum_{k=1}^{m}\|(S_\alpha-s_\alpha)f_k\|^2\sum_{\alpha=1}^{n}\sum_{k=1}^{m}\|(T_\alpha-t_\alpha)f_k\|^2. \tag{6.5}$$

Equality in (6.5) *holds iff*

$$\exists\lambda\in\mathbb{R} : (S_\alpha - s_\alpha)f_k = i\lambda(T_\alpha - t_\alpha)f_k,\ \forall\alpha = 1,\dots,n,\ k = 1,\dots,m.$$

If $\|f_k\| = 1$, *then equality in* (6.5) *implies that*

$$s_\alpha = \langle S_\alpha f_k, f_k\rangle, \text{ and } t_\alpha = \langle T_\alpha f_k, f_k\rangle.$$

Proof. The idea to use the Cauchy-Schwartz inequality to combine a set of uncertainty relations into a single one is inspired by a proof of a Heisenberg uncertainty relation in n-D which can be found in [17] Corollary 2.8 see Remark 6.3.11 for details.

We now use this idea to proof the uncertainty relation for sets of self-adjoint operators applied to vector valued functions.

By theorem Theorem 6.2.1 we know that

$$\big|\langle[S_\alpha,T_\alpha]f_k,f_k\rangle\big| \le 2\|(S_\alpha-s_\alpha)f_k\|\|(T_\alpha-t_\alpha)f_k\|,\ \forall\alpha=1,\dots,n;\ k=1,\dots,m.$$

Summation over α and k and application of the Cauchy-Schwarz inequality for $\mathbb{R}^n$ yields

$$\begin{aligned}\sum_{\alpha=1}^{n}\sum_{k=1}^{m}\big|\langle[S_\alpha,T_\alpha]f_k,f_k\rangle\big\| &\le 2\sum_{\alpha=1}^{n}\sum_{k=1}^{n}\|(S_\alpha-s_\alpha)f_k\|\|(T_\alpha-t_\alpha)f_k|\\ &\le 2\sqrt{\sum_{\alpha=1}^{n}\sum_{k=1}^{m}\|(S_\alpha-s_\alpha)f_k\|^2}\sqrt{\sum_{\alpha=1}^{n}\sum_{k=1}^{m}\|(T_\alpha-t_\alpha)f_k\|^2}.\end{aligned}$$

Equality holds iff

$$\exists \lambda \in i\mathbb{R} : (S_\alpha - s_\alpha) f_k = \lambda (T_\alpha - t_\alpha) f_k,\ \forall \alpha = 1, \ldots, n;\ k = 1, \ldots, m.$$

In this case it follows that

$$\langle (S_\alpha - s_\alpha) f_k, f_k \rangle = \lambda \langle (T_\alpha - t_\alpha) f_k, f_k \rangle,\ \forall \alpha = 1, \ldots, n;\ k = 1, \ldots, m,$$

and hence $\lambda = 0$.

Let $\|f_k\| = 1$. Then $s_\alpha = \langle S_\alpha f_k, f_k \rangle$ and by the same argument $t_\alpha = \langle T_\alpha f_k, f_k \rangle$.

□

6.3.2 Uncertainty relations in Clifford-Hilbert modules

The infinitesimal operators of the Heisenberg group computed in Proposition 6.1.2 are not self adjoint. To apply the uncertainty relation we multiplied by i to get self adjoint operators. In a Clifford algebra setting we interpret the basis vectors as hypercomplex elements of an algebra, hence the operators $S = \sum_\alpha e_\alpha s_\alpha$ and $T = \sum_\alpha e_\alpha t_\alpha$ are self adjoint.

To apply an uncertainty relation to these operators we have to extend Theorem 6.2.1 to the setting of Clifford-Hilbert modules which we introduced in chapter 3.2. By this approach we will state a novel uncertainty relation for Clifford-Hilbert modules.

Theorem 6.3.3 (Uncertainty relation in Cliffor-Hilbert modules)
LET *H be a complex Hilbert space. Furthermore let S, T be densely defined self-adjoint left-linear operators on H_n.*

$$S : H_n \supseteq \mathrm{Dom}(S) \to H_n$$
$$T : H_n \supseteq \mathrm{Dom}(T) \to H_n.$$

The **commutator** $[S, T]$ *is given by*

$$[S, T] := ST - TS.$$

THEN

(i)

$$\left| \big([S, T] f, f \big) \right|^2 \leq 4 \|Sf\|^2 \|Tf\|^2. \tag{6.6}$$

(ii) Equality in (6.6) *holds iff*

$$\forall \boldsymbol{\alpha} \in \mathscr{O}_n \exists \lambda \in \mathbb{R}_0^+ : \langle Sf \rangle_{\boldsymbol{\alpha}} = i\lambda \langle Tf \rangle_{\boldsymbol{\alpha}}. \tag{6.7}$$

And furthermore

$$\|\langle Tf \rangle_{\boldsymbol{\alpha}}\| \|\langle Sf \rangle_{\boldsymbol{\beta}}\| = \|\langle Tf \rangle_{\boldsymbol{\beta}}\| \|\langle Sf \rangle_{\boldsymbol{\alpha}}\|, \forall \boldsymbol{\alpha}, \boldsymbol{\beta} \in \mathscr{O}_n. \tag{6.8}$$

(Recall (3.1)*: $\mathscr{O}_n := \big\{ \boldsymbol{\alpha} \in \mathscr{P}(\{1, \ldots, n\}) : 1 \leq \alpha_1 < \alpha_2 < \ldots < \alpha_\nu \leq n;\ \nu \in \{1, \ldots, n\} \cup \{0\} \big\}$.)*

Proof. ad(i)

$$\begin{aligned}([S,T]f,f) &= (Tf,Sf)-(Sf,Tf)\\ &= \sum_{\boldsymbol{\alpha}\in\mathscr{O}_n}\Big\langle\langle Tf\rangle_{\boldsymbol{\alpha}},\langle Sf\rangle_{\boldsymbol{\alpha}}\Big\rangle_H-\Big\langle\langle Sf\rangle_{\boldsymbol{\alpha}},\langle Tf\rangle_{\boldsymbol{\alpha}}\Big\rangle_H\\ &=2\sum_{\boldsymbol{\alpha}\in\mathscr{O}_n}\Im\big(\langle\langle Tf\rangle_{\boldsymbol{\alpha}},\langle Sf\rangle_{\boldsymbol{\alpha}}\rangle\big).\end{aligned}$$

Hence

$$\begin{aligned}|([S,T]f,f)|^2 &=\Big(2\Big|\sum_{\boldsymbol{\alpha}\in\mathscr{O}_n}\Im\big(\langle\langle Tf\rangle_{\boldsymbol{\alpha}},\langle Sf\rangle_{\boldsymbol{\alpha}}\rangle\big)\Big|\Big)^2\\ &\le 4\Big(\sum_{\boldsymbol{\alpha}\in\mathscr{O}_n}\big|\Im\big(\langle\langle Tf\rangle_{\boldsymbol{\alpha}},\langle Sf\rangle_{\boldsymbol{\alpha}}\rangle\big)\big|\Big)\Big(\sum_{\boldsymbol{\beta}\in\mathscr{O}_n}\big|\Im\big(\langle\langle Tf\rangle_{\boldsymbol{\beta}},\langle Sf\rangle_{\boldsymbol{\beta}}\rangle\big)\big|\Big) \qquad (6.9)\\ &\le 4\sum_{\boldsymbol{\alpha},\boldsymbol{\beta}\in\mathscr{O}_n}\|\langle Tf\rangle_{\boldsymbol{\alpha}}\|\|\langle Sf\rangle_{\boldsymbol{\alpha}}\|\|\langle Tf\rangle_{\boldsymbol{\beta}}\|\|\langle Sf\rangle_{\boldsymbol{\beta}}\| \qquad (6.10)\\ &=4\Big(2\sum_{\boldsymbol{\alpha}<\boldsymbol{\beta}\in\mathscr{O}_n}\|\langle Tf\rangle_{\boldsymbol{\alpha}}\|\|\langle Sf\rangle_{\boldsymbol{\beta}}\|\|\langle Tf\rangle_{\boldsymbol{\beta}}\|\|\langle Sf\rangle_{\boldsymbol{\alpha}}\|\\ &\quad+\sum_{\boldsymbol{\alpha}\in\mathscr{O}_n}\|\langle Tf\rangle_{\boldsymbol{\alpha}}\|^2\|\langle Sf\rangle_{\boldsymbol{\alpha}}\|^2\Big) \qquad (6.11)\\ &\le 4\sum_{\boldsymbol{\alpha},\boldsymbol{\beta}\in\mathscr{O}_n}\|\langle Tf\rangle_{\boldsymbol{\alpha}}\|^2\|\langle Sf\rangle_{\boldsymbol{\beta}}\|^2 \qquad (6.12)\\ &=4\|Tf\|^2\|Sf\|^2,\end{aligned}$$

where (6.10) holds by the Cauchy-Schwartz inequality and (6.12) holds by the inequality $2ab\le a^2+b^2$, $\forall a,b\in\mathbb{R}$. The order of O_n in (6.11) may be chosen arbitrarily, an example would be the lexicographical order.

ad(ii) For equality in (6.6) to hold we require equality in (6.10), (6.11) and (6.12).

(6.10) holds with equality iff $\forall\boldsymbol{\alpha}\in\mathscr{O}_n\,\exists\lambda\in\mathbb{R}:\langle(S-s)f\rangle_{\boldsymbol{\alpha}}=i\lambda\langle(T-t)f\rangle_{\boldsymbol{\alpha}}$. Now (6.9) holds with equality iff $\lambda\ge 0$.

Finally (6.12) holds with equality, iff (6.8) is fulfilled.

□

Remark 6.3.4

The right hand sides of Theorem 6.3.2,

$$4\sum_{j=1}^{n}\sum_{k=1}^{m}\|(S_j-s_j)f_k\|^2\sum_{j=1}^{n}\sum_{k=1}^{m}\|(T_j-t_j)f_k\|^2$$

and Theorem 6.3.3,

$$4\|Sf\|^2\|Tf\|^2=\sum_{\boldsymbol{\alpha}\in\mathscr{O}_n}\|\langle Sf\rangle_{\boldsymbol{\alpha}}\|^2\sum_{\boldsymbol{\beta}\in\mathscr{O}_n}\|\langle Tf\rangle_{\boldsymbol{\beta}}\|^2$$

coincide if the vector part of the operators S,T or the function f in Theorem 6.3.3 vanishes. If the vector part of the operator vanishes, then

$$\langle Sf\rangle_{\boldsymbol{\alpha}}=S_0f_{\boldsymbol{\alpha}}$$

which, for $n=1$, yields Theorem 6.3.2.

If the vector part of the function f vanishes, then

$$\langle Sf\rangle_{\boldsymbol{\alpha}} = S_{\boldsymbol{\alpha}} f_0$$

which, for $m = 1$, yields Theorem 6.3.2.

The most important point in which the left hand sides differ is that the left hand side of Theorem 6.3.2 $\left(\sum_{j=1}^n \sum_{k=1}^m \left|\langle [S_j, T_j] f_k, f_k\rangle\right|\right)^2$ is translation invariant, while the left hand side of Theorem 6.3.3 $\left|\left([S,T]f,f\right)\right|^2$ is not. This can be amended with additional prerequisites - this is the topic of the following section.

Translation invariant uncertainty relations

In the following three corollaries we will address the problem that in general

$$[S-s, T-t] \neq [S,T],$$

that is, the uncertainty principle Theorem 6.3.3 is not translation invariant.

Corollary 6.3.8 and Theorem 6.3.3 are special cases of Theorem 6.3.2, for $n = 1$, respectively, $m = 1$. Corollary 6.3.6 is a special case of Theorem 6.3.2 if $S_\alpha f_\beta = S_\beta f_\alpha$ and $T_\alpha f_\beta = T_\beta f_\alpha$, $\forall \alpha, \beta \in \{1, \ldots, n\}$.

This section shows the difficulties in deriving a translation invariant version of Theorem 6.3.3 – whenever applicable Theorem 6.3.2 is stronger and easier to use than Theorem 6.3.3. The proper domain of application for Theorem 6.3.3 would be operators S, T and functions f which use the full range of indices $\mathscr{O}_n$.

Remark 6.3.5 (Linear operators in Clifford-Hilbert modules)
We have seen in Theorem 3.2.7 (viii) that an operator O on the Clifford-Hilbert module has the form

$$O = \sum_{\boldsymbol{\alpha}\in\mathscr{O}_n} e_{\boldsymbol{\alpha}} O_{\boldsymbol{\alpha}},$$

where $O_{\boldsymbol{\alpha}} : H \to H$.

Note that in order to be left-linear the operators are always applied from the right, i.e.,

$$Of = \sum_{\boldsymbol{\alpha},\boldsymbol{\beta}\in\mathscr{O}_n} e_{\boldsymbol{\alpha}} e_{\boldsymbol{\beta}} O_{\boldsymbol{\beta}}(f_{\boldsymbol{\alpha}}).$$

This means that the linear operator $c\,\mathrm{Id}$ is a multiplication by c from the right rather than the action of $\mathbb{C}_n$ on H_N which is multiplication from the left by c.

If O is not defined on the whole Hilbert-Clifford module denote

$$\mathrm{Dom}(O) = \bigcap_{\boldsymbol{\alpha}\in\mathscr{O}_n} \mathrm{Dom}(O_{\boldsymbol{\alpha}})$$

and

$$\mathrm{Dom}_n(O) := \sum_{\alpha\in\mathscr{O}_n} e_{\boldsymbol{\alpha}} \mathrm{Dom}(O).$$

Of is well defined for some $f \in L_n^2(\mathbb{R}^n)$ iff $f \in \mathrm{Dom}_n(O)$, i.e.

$$f_{\boldsymbol{\alpha}} \in \mathrm{Dom}(O),\ \forall \boldsymbol{\alpha} \in \mathscr{O}_n.$$

Corollary 6.3.6
LET S,T *be as in Theorem 6.3.3 and such that* $S=\sum_{\alpha=1}^{n} e_\alpha S_\alpha$ *and* $T=\sum_{\alpha=1}^{n} e_\alpha T_\alpha$. *Furthermore, let* $s,t\in\mathbb{R}^n\subset\mathbb{R}_n$ *be two vectors and let* $f=\sum_{\alpha=1}^{n} e_\alpha f_\alpha \in F\subseteq H_n\cap \mathrm{Dom}([S,T])$. *Finally, let for all* $\alpha,\beta\in\{1,\dots,n\}$,

$$\begin{aligned}&\Big\langle\big((S_\alpha-s_\alpha)(T_\beta-t_\beta)+(T_\beta-t_\beta)(S_\alpha-s_\alpha)\big)f_\alpha,f_\beta\Big\rangle\\ =&\Big\langle\big((S_\alpha-s_\alpha)(T_\beta-t_\beta)+(T_\beta-t_\beta)(S_\alpha-s_\alpha)\big)f_\beta,f_\alpha\Big\rangle.\end{aligned}\tag{6.13}$$

THEN

$$\big|([S-s,T-t]f,f)\big|=\Big|\sum_{\alpha,\beta=1}^{n}\langle[S_\alpha,T_\alpha]f_\beta,f_\beta\rangle\Big|,$$

and hence

$$\Big|\sum_{\alpha,\beta=1}^{n}\langle[S_\alpha,T_\alpha]f_\beta,f_\beta\rangle\Big|^2\le 4\|(S-s)f\|^2\|(T-t)f\|^2.$$

Remark 6.3.7
(6.13) *is a technical condition deduced from the calculation given in the following proof. It is fulfilled for example if* $f_\alpha=f_\beta$, $\forall\alpha,\beta=1,\dots,n$.

Proof.

$$\begin{aligned}([S-s,T-t]f,f)=&\sum_{\substack{\alpha,\beta,\gamma,\delta=1,\dots,n:\\ e_\gamma e_\alpha e_\beta=\pm e_\delta}}\langle(S_\alpha-s_\alpha)(T_\beta-t_\beta)f_\gamma,f_\delta\rangle e_\gamma\overline{e_\delta}e_\beta e_\alpha\\ &-\langle(T_\beta-t_\beta)(S_\alpha-s_\alpha)f_\gamma f_\delta\rangle e_\gamma\overline{e_\delta}e_\alpha e_\beta\\ =&\sum_{\alpha,\beta=1}^{n}\langle[S_\alpha-s_\alpha,T_\alpha-t_\alpha]f_\beta,f_\beta\rangle\underbrace{e_\beta\overline{e_\beta}e_\alpha e_\alpha}_{=-1}\\ &+\sum_{\substack{\alpha,\beta=1,\dots,n\\ \alpha\neq\beta}}\Bigg(\Big\langle\big((S_\alpha-s_\alpha)(T_\beta-t_\beta)+(T_\beta-t_\beta)(S_\alpha-s_\alpha)\big)f_\alpha,f_\beta\Big\rangle\underbrace{e_\alpha\overline{e_\beta}e_\beta e_\alpha}_{=-1}\\ &\quad+\Big\langle\big((S_\alpha-s_\alpha)(T_\beta-t_\beta)+(T_\beta-t_\beta)(S_\alpha-s_\alpha)\big)f_\beta,f_\alpha\Big\rangle\underbrace{e_\beta\overline{e_\alpha}e_\beta e_\alpha}_{=1}\Bigg)\\ =&-\sum_{\alpha,\beta=1}^{n}\langle[S_\alpha-s_\alpha,T_\alpha-t_\alpha]f_\beta,f_\beta\rangle\\ =&-\sum_{\alpha,\beta=1}^{n}\langle[S_\alpha,T_\alpha]f_\beta,f_\beta\rangle.\end{aligned}$$

□

Corollary 6.3.8
LET S,T *be densely defined self adjoint operators*

$$S:H_n\supseteq\mathrm{Dom}_n(S)\to H_n,$$
$$T:H_n\supseteq\mathrm{Dom}_n(T)\to H_n.$$

Let $f=f_0e_0\in H_n$, *where* $f_0\in H$. *Furthermore let* $s,t\in\mathbb{C}_n$ *be a pair of Clifford algebra valued numbers such that* $s\,\mathrm{Id}$ *and* $t\,\mathrm{Id}$ *are self adjoint, as in Example 3.2.2 on page 57.*

THEN

(i)

$$\left|\left([S,T]f,f\right)\right|^2 = \Big|\sum_{\boldsymbol{\alpha}\in\mathscr{O}_n} \langle [S_{\boldsymbol{\alpha}}, T_{\boldsymbol{\alpha}}] f_0, f_0\rangle\Big|^2 \leq 4\|(S-s)f\|^2\|(T-t)f\|^2. \tag{6.14}$$

(ii) Equality in (6.14) *holds iff*

$$\exists \lambda \in \mathbb{R}_0^+ : (S-s)f = i\lambda(T-t)f, \tag{6.15}$$

and furthermore

$$\|(T-t)_{\boldsymbol{\alpha}} f_0\|\|(S-s)_{\boldsymbol{\beta}} f_0\| = \|(T-t)_{\boldsymbol{\beta}} f_0\|\|(S-s)_{\boldsymbol{\alpha}} f_0\|, \ \forall \boldsymbol{\alpha},\boldsymbol{\beta} \in \mathscr{O}_n. \tag{6.16}$$

(iii) If $\|f\| = 1$*, then equality in* (6.14) *implies that*

$$s = \sum_{\boldsymbol{\alpha}\in\mathscr{O}_n}^{n} e_{\boldsymbol{\alpha}} \langle S_{\boldsymbol{\alpha}} f_0, f_0\rangle, \text{ and } t = \sum_{\boldsymbol{\alpha}\in\mathscr{O}_n} e_{\boldsymbol{\alpha}} \langle T_{\boldsymbol{\alpha}} f_0, f_0\rangle.$$

Proof. Let $S = \sum_{\boldsymbol{\alpha}\in\mathscr{O}_n} e_{\boldsymbol{\alpha}} S_{\boldsymbol{\alpha}}$, $T = \sum_{\boldsymbol{\alpha}\in\mathscr{O}_n} e_{\boldsymbol{\alpha}} T_{\boldsymbol{\alpha}}$, where $S_{\boldsymbol{\alpha}} : \mathrm{Dom}(S) \to H$ and $T_{\boldsymbol{\alpha}} : \mathrm{Dom}(S) \to H$. Furthermore let $s = \sum_{\boldsymbol{\alpha}\in\mathscr{O}_n} e_{\boldsymbol{\alpha}} s_{\boldsymbol{\alpha}}$, $t = \sum_{\boldsymbol{\alpha}\in\mathscr{O}_n} e_{\boldsymbol{\alpha}} t_{\boldsymbol{\alpha}}$, where $s_{\boldsymbol{\alpha}}, t_{\boldsymbol{\alpha}} \in \mathbb{C}$.

ad(i) We only need to show that $\left([(S-s),(T-t)]f,f\right) = \left([S,T]f,f\right)$. The rest follows by Theorem 6.3.3(i).

Let $s,t \in \mathbb{C}_n$. Then

$$[(S-s),(T-t)] = (ST-TS) - (sT-Ts) - (St-tS) + (st-ts).$$

This is in general certainly non zero, since the Clifford algebra is not commutative. Let us consider

$$\begin{aligned}\left([S-s,T-t]f,f\right) &= \left\langle \langle [(S-s),(T-t)]\rangle_0 f_0, f_0\right\rangle_H \\ &= \left\langle \langle (ST-TS)-(sT-Ts)-(St-tS)+(st-ts)\rangle_0 f_0, f_0\right\rangle_H.\end{aligned}$$

Hence, if the identity

$$\langle (st-ts)-(sT-Ts)-(St-tS)\rangle_0 = 0 \tag{6.17}$$

holds, it is true that

$$\left([S-s,T-t]f,f\right) = \left([S,T]f,f\right), \ \forall f \in H_n : f_{\boldsymbol{\alpha}} = 0, \ \forall \boldsymbol{\alpha} \neq 0.$$

Indeed, since $s_{\boldsymbol{\beta}}, t_{\boldsymbol{\beta}} \in \mathbb{C}$ and $S_{\boldsymbol{\beta}}, T_{\boldsymbol{\beta}} \in H$, $\forall \boldsymbol{\beta} \in \mathscr{O}_n$,

$$\langle st-ts\rangle_0 = \sum_{\boldsymbol{\beta}=1}^{n} e_{\boldsymbol{\beta}}^2 (s_{\boldsymbol{\beta}} t_{\boldsymbol{\beta}} - t_{\boldsymbol{\beta}} s_{\boldsymbol{\beta}}) = 0,$$

$$\langle St-tS\rangle_0 = \sum_{\boldsymbol{\beta}=1}^{n} e_{\boldsymbol{\beta}}^2 (S_{\boldsymbol{\beta}} t_{\boldsymbol{\beta}} - t_{\boldsymbol{\beta}} S_{\boldsymbol{\beta}}) = 0,$$

and finally

$$\langle sT-Ts\rangle_0 = \sum_{\boldsymbol{\beta}=1}^{n} e_{\boldsymbol{\beta}}^2 (s_{\boldsymbol{\beta}} T_{\boldsymbol{\beta}} - T_{\boldsymbol{\beta}} s_{\boldsymbol{\beta}}) = 0.$$

Now $(S-s)$ and $(T-t)$ satisfy the conditions for Theorem 6.3.3 which proves (i).

ad(ii) See Theorem 6.3.3(ii).

ad(iii) First note that $\langle\!\langle Sf,f\rangle\!\rangle\mathrm{Id}$ and $\langle\!\langle Tf,f\rangle\!\rangle\mathrm{Id}$ are self adjoint. For this note that by Proposition 3.2.10(c) $\langle S_{\boldsymbol{\alpha}} f_0, f_0\rangle \in \mathbb{R}$, $\forall \boldsymbol{\alpha} : \overline{e_{\boldsymbol{\alpha}}} = e_{\boldsymbol{\alpha}}$ and $\langle S_{\boldsymbol{\alpha}} f_0, f_0\rangle \in i\mathbb{R}$ $\forall \boldsymbol{\alpha} : \overline{e_{\boldsymbol{\alpha}}} = e_{\boldsymbol{\alpha}}$. But this means that $\overline{\langle\!\langle Sf,f\rangle\!\rangle} = \langle\!\langle Sf,f\rangle\!\rangle$ and hence that $\langle\!\langle Sf,f\rangle\!\rangle\mathrm{Id}$ is self adjoint.

Now if we take the inner product of both sides of (6.15) we get

$$\begin{aligned}
&\langle\!\langle (S-s)f,f\rangle\!\rangle = i\lambda\langle\!\langle (T-t)f,f\rangle\!\rangle\\
\Leftrightarrow &\sum_{\boldsymbol{\alpha}\in\mathcal{O}_n} e_{\boldsymbol{\alpha}}\langle\langle S-s\rangle_{\boldsymbol{\alpha}} f_0, f_0\rangle = i\lambda \sum_{\boldsymbol{\alpha}\in\mathcal{O}_n} e_{\boldsymbol{\alpha}}\langle\langle S-s\rangle_{\boldsymbol{\alpha}} f_0, f_0\rangle\\
\Leftrightarrow &\langle\langle S-s\rangle_{\boldsymbol{\alpha}} f_0, f_0\rangle = i\lambda\langle\langle T-t\rangle_{\boldsymbol{\alpha}} f_0, f_0\rangle,\ \forall \boldsymbol{\alpha}\in\mathcal{O}_n\\
\Leftrightarrow &\langle S_{\boldsymbol{\alpha}} f_0, f_0\rangle = \langle s\rangle_{\boldsymbol{\alpha}} \text{ and } \langle T_{\boldsymbol{\alpha}} f_0, f_0\rangle = \langle t\rangle_{\boldsymbol{\alpha}},\ \forall \boldsymbol{\alpha}\in\mathcal{O}_n\\
\Leftrightarrow &s = \langle\!\langle Sf,f\rangle\!\rangle \text{ and } t = \langle\!\langle Tf,f\rangle\!\rangle.
\end{aligned}$$

□

Corollary 6.3.9
LET S_0, T_0 *be densely defined self adjoint operators*

$$\begin{aligned}
S &: H \supseteq \mathrm{Dom}(S) \to H\\
T &: H \supseteq \mathrm{Dom}(T) \to H,
\end{aligned}$$

and let $S = e_0 S_0$, $T = e_0 T_0 : H_n \to H_n$. *Let* $f \in H_n$, *and let* $s, t \in \mathbb{R}$.
THEN

(i)

$$\left|\left([S,T]f,f\right)\right|^2 = \Big|\sum_{\alpha=1}^{n}\langle [S_0,T_0]f_\alpha, f_\alpha\rangle\Big|^2 \le 4\|(S-s)f\|^2\|(T-t)f\|^2. \tag{6.18}$$

(ii) Equality in (6.18) *holds iff*

$$\forall \boldsymbol{\alpha}\in\mathcal{O}_n\, \exists\lambda\in\mathbb{R}_0^+ : (S_0-s_0)f_{\boldsymbol{\alpha}} = i\lambda(T_0-t_0)f_{\boldsymbol{\alpha}}. \tag{6.19}$$

And furthermore for all $\alpha,\beta\in\mathcal{O}_n$

$$\|(T_0-t_0)f_{\boldsymbol{\alpha}}\|\|(S-s)f_{\boldsymbol{\beta}}\| = \|(T-t)f_{\boldsymbol{\beta}}\|\|(S-s)f_{\boldsymbol{\alpha}}\| \tag{6.20}$$

(iii) If $\|f_{\boldsymbol{\alpha}}\| = 1$, $\forall \boldsymbol{\alpha}\in O\subseteq\mathcal{O}_n$, *and* $f_{\boldsymbol{\alpha}} = 0$; $\forall \boldsymbol{\alpha}\notin O$, *then equality in* (6.18) *implies that for all* $\boldsymbol{\alpha}\in O$

$$s_0 = \langle S_0 f_{\boldsymbol{\alpha}}, f_{\boldsymbol{\alpha}}\rangle \text{ and } t_0 = \langle T_0 f_{\boldsymbol{\alpha}}, f_{\boldsymbol{\alpha}}\rangle.$$

Proof.

ad(i) We only need to show that $\left([(S-s),(T-t)]f,f\right) = \left([S,T]f,f\right)$. The rest follows by Theorem 6.3.3(i).

Indeed

$$\begin{aligned}
[S-s,T-t] &= ST - TS - sT + Ts - St + tS + st - ts\\
&= e_0\big(S_0T_0 - T_0S_0 - s_0T_0 + T_0s_0 - S_0t_0 + t_0S_0 + s_0t_0 - t_0s_0\big) = e_0(S_0T_0 - T_0S_0)\\
&= [S,T].
\end{aligned}$$

ad(ii) See Theorem 6.3.3(ii).

ad(iii) Let $\|f\| = 1$. Then equality in (6.18) implies (6.19) and hence that

$$\exists \lambda \in \mathbb{R}_0^+ : (S_0 - s_0) f_{\boldsymbol{\alpha}} = i\lambda (T_0 - t_0) f_{\boldsymbol{\alpha}},\ \forall \boldsymbol{\alpha} \in \mathscr{O}_n.$$

It follows that

$$\langle (S_0 - s_0) f_{\boldsymbol{\alpha}}, f_{\boldsymbol{\alpha}} \rangle = i\lambda \langle (T_0 - t_0) f_{\boldsymbol{\alpha}}, f_{\boldsymbol{\alpha}} \rangle.$$

Since $(S_0 - s_0)$ and $(T_0 - t_0)$ are self adjoint operators, we know that $\lambda = 0$ and for all $\boldsymbol{\alpha} \in \mathscr{O}_n$

$$s_0 \langle f_{\boldsymbol{\alpha}}, f_{\boldsymbol{\alpha}} \rangle = \langle S_0 f_{\boldsymbol{\alpha}}, f_{\boldsymbol{\alpha}} \rangle$$
$$t_0 \langle f_{\boldsymbol{\alpha}}, f_{\boldsymbol{\alpha}} \rangle = \langle T_0 f_{\boldsymbol{\alpha}}, f_{\boldsymbol{\alpha}} \rangle.$$

□

6.3.3 The Weyl-Heisenberg uncertainty

Applying Corollary 6.3.8 to the infinitesimal operators derived from the n-dimensional Weyl-Heisenberg group Definition 6.1.1 by the Dirac operator $D = \sum_{\alpha=1}^n e_\alpha \frac{\partial}{\partial x_\alpha}$ leads to the isotropic uncertainty relation stated below. This theorem is already known ([17] Corollary 2.8). However, we will give a new proof. This proof is based on the fact that the operator D is the square root of the Laplace operator.

While we are not in a setting where applying Theorem 6.3.3 via Corollary 6.3.8 is strictly necessary, the new proof shows the concepts behind the Weyl-Heisenberg uncertainty relation – vector valued infinitesimal operators, which are square roots of the modulus respectively the Laplace operator.

Theorem 6.3.10 (A single Weyl-Heisenberg uncertainty relation for $\mathbb{R}^n$)
LET $f_0 \in L^2(\mathbb{R}^n)$, $l, r \in \mathbb{R}^n$.
THEN *the following inequality holds:*

$$16\pi^2 \int_{\mathbb{R}^n} |x+l|^2 |\widehat{f}(x)|^2 dx \int_{\mathbb{R}^n} |x+r|^2 |f(x)|^2 dx \geq n^2 \|f\|^4.$$

Equality holds if $f(x) = e^{-\frac{\pi}{\lambda}|x|^2}$, *for some* $\lambda \in \mathbb{R}^+$.

Proof. Let $s = \sum_{\alpha=1}^n e_\alpha s_\alpha = 2\pi i \sum_{\alpha=1}^n e_\alpha r_\alpha$ and $t = \sum_{\alpha=1}^n e_\alpha t_\alpha = 2\pi i \sum_{\alpha=1}^n e_\alpha l_\alpha$. Note that $s\mathrm{Id}$ and $t\mathrm{Id}$ are self-adjoint left linear operators on $L_n^2(\mathbb{R}^n)$. (See Example 3.2.2.)
In order to proof the theorem we apply inequality (6.14), which states that

$$4\|(S-s)f\|^2 \|(T-t)f\|^2 \geq \left|\left([S,T]f, f\right)\right|^2,$$

to the function $f = e_0 f_0$ and the operators given for suitable functions in $g \in L^2(\mathbb{R}^n)_n$ by

$$Sg(x) = 2\pi i \sum_{\boldsymbol{\beta} \in \mathscr{O}_n} \sum_{\alpha=1}^n e_{\boldsymbol{\beta}} e_\alpha x_\alpha g_{\boldsymbol{\beta}}(x)$$

and

$$Tg(x) = Dg(x) = \sum_{\boldsymbol{\beta}\in\mathscr{O}_n}\sum_{\alpha=1}^{n} e_{\boldsymbol{\beta}} e_\alpha \frac{\partial}{\partial x_\alpha} g_{\boldsymbol{\beta}}(x),$$

respectively.

The right hand side of (6.14) yields

$$\begin{aligned}
[S,T]f(x) =& STf(x) - TSf(x)\\
=& -2\pi i \sum_{\alpha,\beta=1}^{n} e_\beta e_\alpha x_\alpha \frac{\partial}{\partial x_\beta} f(x) + 2\pi i \sum_{\beta,\alpha=1}^{n} e_\alpha e_\beta \frac{\partial}{\partial x_\beta}(x_\alpha f(x))\\
=& 2\pi i \Big(\sum_{\alpha=1}^{n} x_\alpha \frac{\partial}{\partial x_\alpha} f(x) - \frac{\partial}{\partial x_\alpha}\big(x_\alpha f(x)\big)\\
&+ 2\sum_{\alpha<\beta=1}^{n} e_\alpha e_\beta \big(x_\alpha \frac{\partial}{\partial x_\beta} f(x) + x_\beta \frac{\partial}{\partial x_\alpha} f(x)\big)\Big)\\
=& 2\pi i \Big(-nf(x) + 2\sum_{\alpha<\beta=1}^{n} e_\alpha e_\beta \big(x_\alpha\big(\frac{\partial}{\partial x_\beta} f(x)) + x_\beta(\frac{\partial}{\partial x_\alpha} f(x)\big)\big)\Big).
\end{aligned}$$

Hence

$$\big|\big([S,T]f,f\big)\big|^2 = \big|\langle\langle [S,T]f\rangle_0, f_0\rangle\big|^2 = |2\pi i n|^2 \|f_0\|^4.$$

We now consider the term

$$\|(S-s)f\|^2.$$

Note that $2\pi i x + s$ is a vector and hence an element of the Clifford group and furthermore $2\pi i x + s = \overline{2\pi i x + s}$, whence $(2\pi i x + s)\overline{(2\pi i x + s)} = |2\pi i x + s|^2 = (2\pi i x + s)^2$.

$$\begin{aligned}
\|(S-s)f\|^2 &= \big((S-s)^2 f, f\big)\\
&= \Big\langle \sum_{\alpha=1}^{n}\sum_{\beta=1}^{n} e_\beta e_\alpha \int_{\mathbb{R}^n} (-2\pi i x_\alpha - s_\alpha)(-2\pi i x_\beta - s_\beta) f(x)\overline{f(x)}\,dx \Big\rangle_0\\
&= \int_{\mathbb{R}^n} (-2\pi i x - s)\overline{(-2\pi i x - s)} |f(x)|^2 dx = \int_{\mathbb{R}^n} |-2\pi i x - s|^2 |f(x)|^2 dx\\
&= \int_{\mathbb{R}^n} \sum_{\alpha=1}^{n} (2\pi i x_\alpha + s_\alpha)^2 |f(x)|^2 dx\\
&= 4\pi^2 \int_{\mathbb{R}^n} |x + r|^2 |f(x)|^2 dx.
\end{aligned}$$

Furthermore,

$$\begin{aligned}
\|(T-t)f\|^2 = \|(S-t)\widehat{f}\|^2 &= -\int_{\mathbb{R}^n} \sum_{\alpha=1}^{n} (2\pi i x_\alpha + t_\alpha)^2 |\widehat{f}(x)|^2 dx\\
&= 4\pi^2 \int_{\mathbb{R}^n} |x + l|^2 |\widehat{f}(x)|^2 dx.
\end{aligned}$$

Equality holds if

$$\|\langle (T-t)f\rangle_{\boldsymbol{\alpha}}\| \|\langle (S-s)f\rangle_{\boldsymbol{\beta}}\| = \|\langle (T-t)f\rangle_{\boldsymbol{\beta}}\| \|\langle (S-s)f\rangle_{\boldsymbol{\alpha}}\|,$$

which holds if f is a radial function. Furthermore,

$$\exists\lambda \in \mathbb{R}_0^+ : (S-s)f = i\lambda(T-t)f,$$

which holds if $f(x) = e^{-\frac{\pi}{\lambda}\sum_{\alpha=1}^{n} x_\alpha^2}$, for some $\lambda \in \mathbb{R}^+$.

□

Remark 6.3.11 (A shortcut)
The same inequality could have been derived by Theorem 6.3.2 as well. Indeed the idea of the proof for Theorem 6.3.2 came from a proof of Theorem 6.3.10 which can be found in [17] Corollary 2.8. There Theorem 6.3.10 is proven using Corollary 6.3.1, which yields

$$4\pi\sqrt{\sum_{\alpha=1}^{n}\int_{\mathbb{R}}(x_\alpha + r_\alpha)^2|f(x)|^2dx}\sqrt{\int_{\mathbb{R}}(\xi_\alpha + l_\alpha)^2|\widehat{f}(\xi)|^2d\xi} \geq n\|f\|^2,\ \forall\alpha = 1,\dots,n.$$

Applying the Cauchy-Schwartz inequality yields

$$4\pi\sqrt{\sum_{\alpha=1}^{n}\int_{\mathbb{R}}(x_\alpha + r_\alpha)^2|f(x)|^2dx}\sqrt{\sum_{\beta=1}^{n}\int_{\mathbb{R}}(\xi_\beta + l_\beta)^2|\widehat{f}(\xi)|^2d\xi} \geq n\|f\|^2.$$

6.3.4 The Riesz transform and Heisenberg uncertainty

We will now compute the effect of the Riesz transform on the appropriate version of the uncertainty relation Theorem 6.3.10 derived from Theorem 6.3.2. We cannot study the effects of the Riesz transform directly on Theorem 6.3.10, since the Riesz transform is vector valued.

The essence of the next theorem is that the Heisenberg uncertainty relation is not invariant under the Riesz transform since the operator S does not commute with Riesz transforms.

Theorem 6.3.12 (The Heisenberg uncertainty relation and the Riesz transform)
LET $f \in L^2(\mathbb{R}^n, \mathbb{R})$.
THEN

(i) $16\pi^2\int_{\mathbb{R}^n}|x+l|^2\sum_{\alpha=1}^{n}\left|\widehat{R_\alpha f}(x)\right|^2dx\int_{\mathbb{R}^n}|x+r|^2\sum_{\alpha=1}^{n}|R_\alpha f(x)|^2dx \geq n^2\|f\|^4$;

(ii) $\int_{\mathbb{R}^n}|x+l|^2\sum_{\alpha=1}^{n}|\widehat{R_\alpha f}(x)|^2dx = \int_{\mathbb{R}^n}|x+l|^2|\widehat{f}(x)|^2dx$, *i.e. localisation in the frequency domain is invariant under the Riesz transform.*

(iii) *Let* $\widehat{f}, \widehat{R_\alpha f} \in W^{1,2}(\mathbb{R}^n) := \left\{f \in L^2(\mathbb{R}^n) : \partial^\alpha\widehat{f} \in L^2(\mathbb{R}^n), \forall\alpha \in \{1,\dots,n\}\right\}$. *(Sufficient conditions for this to hold will be given in Theorem 5.4.3.)*

Then $\int_{\mathbb{R}^n}|x+r|^2\sum_{\alpha=1}^{n}|R_\alpha f(x)|^2dx = \int_{\mathbb{R}^n}|x+r|^2|f(x)|^2dx + \frac{(n-1)}{4\pi^2}\|\frac{1}{|\cdot|}\widehat{f}\|^2$, *i.e., localization in the space domain is not invariant under the Riesz transform.*

Proof.

$ad(i)$
$$\sum_{j,\alpha=1}^{n}\left|\left\langle [S_j,T_j]R_\alpha f, R_\alpha f\right\rangle\right| \geq n\sum_{\alpha=1}^{n}\left\|\widehat{R_\alpha f}\right\|^2 = n\int_{\mathbb{R}^n}\sum_{\alpha=1}^{n}\frac{x_\alpha^2}{|x|^2}\left|\widehat{f}(x)\right|^2 dx$$
$$= n\|f\|^2.$$

$ad(ii)$
$$\sum_{j,\alpha=1}^{n}\int_{\mathbb{R}^n}(x_j+l_j)\frac{ix_\alpha}{|x|}\widehat{f}(x)\overline{(x_j+l_j)\frac{ix_\alpha}{|x|}\widehat{f}(x)}dx$$
$$=\sum_{j=1}^{n}\int_{\mathbb{R}^n}\sum_{\alpha=1}^{n}\frac{x_\alpha^2}{|x|^2}(x_j+l_j)\widehat{f}(x)\overline{(x_j+l_j)\widehat{f}(x)}dx$$
$$=\int_{\mathbb{R}^n}|x+l|^2|\widehat{f}(x)|^2dx.$$

$ad(iii)$

The additional condition makes sure that $S_j f\in L^2(\mathbb{R}^n)$ and $S_jR_\alpha f\in L^2(\mathbb{R}^n)$.

$$\begin{aligned}
&\sum_{j,\alpha=1}^{n}\int_{\mathbb{R}^n}(x_j+r_j)R_\alpha f(x)\overline{(x_j+r_j)R_\alpha f(x)}dx\\
&=\frac{1}{4\pi^2}\sum_{j,\alpha=1}^{n}\int_{\mathbb{R}^n}\Big(\frac{\partial}{\partial x_j}+2\pi r_j\Big)\frac{ix_\alpha}{|x|}\widehat{f}(x)\overline{\Big(\frac{\partial}{\partial x_j}+2\pi r_j\Big)\frac{ix_\alpha}{|x|}\widehat{f}(x)}dx\\
&=\frac{1}{4\pi^2}\sum_{j,\alpha=1}^{n}\int_{\mathbb{R}^n}\Big(\frac{ix_\alpha}{|x|}\Big(\frac{\partial}{\partial x_j}+2\pi r_j\Big)-\frac{ix_jx_\alpha}{|x|^3}+\frac{i\delta(j,\alpha)}{|x|}\Big)\widehat{f}(x)\\
&\qquad\times\overline{\Big(\frac{ix_\alpha}{|x|}\Big(\frac{\partial}{\partial x_j}+2\pi r_j\Big)-\frac{ix_jx_\alpha}{|x|^3}+\frac{i\delta(j,\alpha)}{|x|}\Big)\widehat{f}(x)}dx\\
&=: I=\int_{\mathbb{R}^n}|x+r|^2|f(x)|^2dx+\frac{(n-1)}{4\pi^2}\left\||\cdot|^{-1}\widehat{f}\right\|^2.
\end{aligned}$$

To see this we compute I as

$$\begin{aligned}
I&=\frac{1}{4\pi^2}\sum_{j,\alpha=1}^{n}\int_{\mathbb{R}^n}\Big(\frac{ix_\alpha}{|x|}\Big(\frac{\partial}{\partial x_j}+2\pi r_j\Big)-\frac{ix_jx_\alpha}{|x|^3}+\frac{i\delta(j,\alpha)}{|x|}\Big)\widehat{f}(x)\\
&\qquad\times\overline{\Big(\frac{ix_\alpha}{|x|}\Big(\frac{\partial}{\partial x_j}+2\pi r_j\Big)-\frac{ix_jx_\alpha}{|x|^3}+\frac{i\delta(j,\alpha)}{|x|}\Big)\widehat{f}(x)}dx\\
&=\int_{\mathbb{R}^n}|x+r|^2|f(x)|^2dx+\frac{1}{4\pi^2}\Big(\underbrace{\sum_{j,\alpha=1}^{n}\int_{\mathbb{R}^n}\frac{ix_jx_\alpha}{|x|^3}\widehat{f}(x)\overline{\frac{ix_jx_\alpha}{|x|^3}\widehat{f}(x)}dx}_{(1)}\\
&\quad+\underbrace{\sum_{j,\alpha=1}^{n}\int_{\mathbb{R}^n}\frac{i\delta(j,\alpha)}{|x|}\widehat{f}(x)\overline{\frac{i\delta(j,\alpha)}{|x|}\widehat{f}(x)}dx}_{(2)}+\underbrace{2\sum_{j,\alpha}\int_{\mathbb{R}^n}\frac{ix_\alpha}{|x|}\Big(\frac{\partial}{\partial x_j}+2\pi r_j\Big)\widehat{f}(x)\overline{\frac{i\delta(j,\alpha)}{|x|}\widehat{f}(x)}dx}_{(3)}\\
&\quad\underbrace{-2\sum_{j,\alpha}\int_{\mathbb{R}^n}\frac{i\delta(j,\alpha)}{|x|}\widehat{f}(x)\overline{\frac{ix_jx_\alpha}{|x|^3}\widehat{f}(x)}dx}_{(4)}\underbrace{-2\sum_{j,\alpha}\int_{\mathbb{R}^n}\frac{ix_\alpha}{|x|}\Big(\frac{\partial}{\partial x_j}+2\pi r_j\Big)\widehat{f}(x)\overline{\frac{ix_jx_\alpha}{|x|^3}\widehat{f}(x)}dx}_{(5)}\Big).
\end{aligned}$$

We continue to compute the terms (1)–(5)

$$(1) = \sum_{j,\alpha=1}^{n} \int_{\mathbb{R}^n} \frac{ix_j x_\alpha}{|x|^3} \widehat{f}(x) \overline{\frac{ix_j x_\alpha}{|x|^3} \widehat{f}(x)} dx = \sum_{j,\alpha} \int_{\mathbb{R}^n} \frac{x_j^2}{|x|^2} \frac{x_\alpha^2}{|x|^2} \frac{1}{|x|^2} |\widehat{f}(x)|^2 dx = \int_{\mathbb{R}^n} \frac{1}{|x|^2} |\widehat{f}(x)|^2 dx.$$

$$(2) = \sum_{j,\alpha=1}^{n} \int_{\mathbb{R}^n} \frac{i\delta(j,\alpha)}{|x|} \widehat{f}(x) \overline{\frac{i\delta(j,\alpha)}{|x|} \widehat{f}(x)} dx = n \int_{\mathbb{R}^n} \frac{1}{|x|^2} |\widehat{f}(x)|^2 dx.$$

$$(3) = 2\sum_{j,\alpha} \int_{\mathbb{R}^n} \frac{ix_\alpha}{|x|} \Big(\frac{\partial}{\partial x_j} + 2\pi r_j \Big) \widehat{f}(x) \overline{\frac{i\delta(j,\alpha)}{|x|} \widehat{f}(x)} dx = 2\sum_{\alpha} \int_{\mathbb{R}^n} \frac{ix_\alpha}{|x|^2} \Big(\frac{\partial}{\partial x_\alpha} + 2\pi r_j \Big) \widehat{f}(x) \overline{\widehat{f}(x)} dx.$$

$$(4) = 2\sum_{j,\alpha} \int_{\mathbb{R}^n} \frac{i\delta(j,\alpha)}{|x|} \widehat{f}(x) \overline{\frac{ix_j x_\alpha}{|x|^3} \widehat{f}(x)} dx = 2\sum_{\alpha} \int_{\mathbb{R}^n} \frac{x_\alpha^2}{|x|^2} \frac{1}{|x|^2} |\widehat{f}(x)|^2 dx = 2\int_{\mathbb{R}^n} \frac{1}{|x|^2} |\widehat{f}(x)|^2 dx.$$

$$(5) = 2\sum_{j,\alpha} \int_{\mathbb{R}^n} \frac{ix_\alpha}{|x|} \Big(\frac{\partial}{\partial x_j} + 2\pi r_j \Big) \widehat{f}(x) \overline{\frac{ix_j x_\alpha}{|x|^3} \widehat{f}(x)} dx = 2\sum_{j} \int_{\mathbb{R}^n} \frac{ix_j}{|x|^2} \Big(\frac{\partial}{\partial x_j} + 2\pi r_j \Big) \widehat{f}(x) \overline{\widehat{f}(x)} dx.$$

Whence (1)+(2)-(4)=$\frac{(n-1)}{4\pi^2} \| \frac{1}{|\cdot|} \widehat{f} \|^2$ and (3)-(5)= 0 and hence

$$I = \int_{\mathbb{R}^n} |x + r|^2 |f(x)|^2 dx + \frac{(n-1)}{4\pi^2} \Big\| \frac{1}{|\cdot|} \widehat{f} \Big\|^2.$$

□

6.3.5 The affine uncertainty

We now consider the uncertainty relation for the infinitesimal operators of the affine group. The affine group $\mathfrak{A}$ is defined in one dimension by

$$\mathfrak{A} = \{(a,b) \in \mathbb{R}^2 : a \neq 0\}$$

with group law

$$(a,b) \cdot (c,d) = (ac, ad + b).$$

As an extension of the affine group of $\mathbb{R}^n$, $n \geq 2$ we consider the group of translations and anisotropic dilations. (Note that there are alternative choices for an affine group in higher dimensions. See for example [3].)

The representation of this group on $L^2(\mathbb{R}^n)$ is given by the operator

$$U_{a,b} f(x) = \prod_{k=1}^{n} |a_k|^{\frac{1}{2}} f\left(\begin{pmatrix} \frac{x_1 - b_1}{a_1} \\ \vdots \\ \frac{x_n - b_n}{a_n} \end{pmatrix} \right).$$

The infinitesimal operators are

$$\begin{aligned}
Af(x) &:= \sum_{\alpha=1}^{n} e_\alpha \frac{\partial}{\partial a_\alpha} U_{a,b} f(x)\Big|_{(a,b)=(1,0)} \\
&= \sum_{\alpha=1}^{n} e_\alpha \Big(\prod_{l=1,\dots,n} |a^l|^{-\frac{1}{2}} \frac{-a_\alpha}{2|a_\alpha|^{\frac{3}{2}}} f\Big(\begin{pmatrix} \frac{x_1-b_1}{a_1} \\ \vdots \\ \frac{x_n-b_n}{a_n} \end{pmatrix} \Big) - \frac{x_\alpha}{a_\alpha^2} \prod_{l=1}^{n} |a_l|^{\frac{1}{2}} \frac{\partial}{\partial x_\alpha} f\Big(\begin{pmatrix} \frac{x_1-b_1}{a_1} \\ \vdots \\ \frac{x_n-b_n}{a_n} \end{pmatrix} \Big)\Big) \\
&= \sum_{\alpha=1}^{n} e_\alpha \Big(-\frac{1}{2} f(x) - x_\alpha \frac{\partial}{\partial x_\alpha} f(x) \Big) \\
Bf(x) &:= \sum_{\alpha=1}^{n} e_\alpha \frac{\partial}{\partial b_\alpha} U_{a,b} f(x)\Big|_{(a,b)=(1,0)} \\
&= \prod_{k=1}^{n} |a_k|^{\frac{1}{2}} \sum_{\alpha=1}^{n} e_\alpha \frac{\partial}{\partial x_\alpha} f\Big(\begin{pmatrix} \frac{x_1-b_1}{a_1} \\ \vdots \\ \frac{x_n-b_n}{a_n} \end{pmatrix} \Big) \frac{-1}{a_\alpha} \\
&= -\sum_{\alpha=1}^{n} e_\alpha \frac{\partial}{\partial x_\alpha} f(x).
\end{aligned}$$

The following theorem is the corresponding uncertainty relation.

Theorem 6.3.13 (An affine uncertainty relation)
LET $f \in W^{1,2}(\mathbb{R}^n)$.
THEN

$$4 \Big\| \sum_{\alpha=1}^{n} \partial_\alpha f \Big\|^2 \Big\| \frac{n}{2} f + \sum_{\alpha=1}^{n} \cdot\ {}_\alpha \partial_\alpha f \Big\|^2 \geq \Big| \Big\langle \frac{1}{2} \sum_{\alpha=1}^{n} \partial_\alpha f, f \Big\rangle \Big|^2, \tag{6.21}$$

where $\partial_\alpha f(x) := \frac{\partial}{\partial x_\alpha} f(x)$, $\forall x \in \mathbb{R}^n$.

Proof. We need to compute $[A, B]$, so we start by considering

$$\begin{aligned}
ABe_0 f(x) &= \sum_{\alpha,\beta=1}^{n} e_\beta e_\alpha \frac{\partial}{\partial x_\alpha} (-\frac{1}{2} - x_\beta \frac{\partial}{\partial x_\beta}) f(x) \\
&= \sum_{\alpha=1}^{n} \frac{\partial}{\partial x_\alpha} (\frac{1}{2} + \frac{\partial}{\partial x_\alpha} x_\alpha + \frac{1}{2}) + \sum_{\alpha,\beta=1}^{n} e_\alpha e_\beta \frac{\partial}{\partial x_\alpha} (\frac{1}{2} + x_\beta \frac{\partial}{\partial x_\beta}) f(x)
\end{aligned}$$

and continue with

$$\begin{aligned}
BAe_0 f(x) &= \sum_{\alpha,\beta=1}^{n} e_\alpha e_\beta (-\frac{1}{2} - x_\alpha \frac{\partial}{\partial x_\alpha}) \frac{\partial}{\partial x_\alpha} f(x) \\
&= \sum_{\alpha=1}^{n} (\frac{1}{2} + x_\alpha \frac{\partial}{\partial x_\alpha}) \frac{\partial}{\partial x_\alpha} f(x) - \sum_{\alpha,\beta=1}^{n} e_\alpha e_\beta \frac{\partial}{\partial x_\alpha} (\frac{1}{2} + x_\beta \frac{\partial}{\partial x_\beta}) f(x).
\end{aligned}$$

Consequently

$$[A, B] e_0 f(x) = (AB - BA) f(x) = \frac{1}{2} \frac{\partial}{\partial x_\alpha} f(x) + 2 \sum_{\alpha,\beta=1}^{n} e_\alpha e_\beta \frac{\partial}{\partial x_\alpha} (\frac{1}{2} + x_\beta \frac{\partial}{\partial x_\beta}) f(x).$$

Hence $([A,B]e_0 f, e_0 f) = \langle \frac{1}{2}\sum_{\alpha=1}^n \frac{\partial}{\partial \cdot_\alpha} f, f\rangle$. Applying (6.6) yields

$$4\Big\| \sum_{\alpha=1}^n \frac{\partial}{\partial \cdot_\alpha}\Big\|^2 \Big\| \frac{n}{2} + \sum_{\alpha=1}^n \cdot_\alpha \frac{\partial}{\partial \cdot_\alpha}\Big\|^2 \leq \Big|\langle \frac{1}{2}\sum_{\alpha=1}^n \frac{\partial}{\partial \cdot_\alpha} f, f\rangle\Big|^2. \tag{6.22}$$

□

6.4 Discussion and comparison to the literature

The first new result in this chapter is Theorem 6.2.2. It states that under the assumption of one vanishing moment the Heisenberg uncertainty relation is invariant under the Hilbert transform. Of course the question arises if a the same holds for the Riesz transform. Since the Riesz transform R is unitary (see Theorem 4.1.4.) whereas the partial Riesz transforms R_α are not it seems reasonable to consider a single uncertainty relation for the Riesz transform rather than a set of uncertainty relations for the partial Riesz transforms. This requires a new kind of uncertainty relation for vector valued functions.

We consider two approaches to such uncertainty relations for vector valued functions - a classical approach Theorem 6.3.2 based on the Cauchy-Schwartz inequality and a second approach on Hilbert-Clifford modules Theorem 6.3.3. The second approach is not translation invariant thus we give three special cases which are translation invariant - Corollary 6.3.6, Corollary 6.3.8 and Corollary 6.3.9. Theorem 6.3.2 is easier to use and gives stronger inequalities than Theorem 6.3.3 and its special cases. However Theorem 6.3.3 is more general.

Theorem 6.3.10 is an example of a single uncertainty relation for operators in higher dimensions known in the literature (see [17] Corollary 2.8.) - however we present the first systematic approach. In [11] an uncertainty relation for vector-valued operators is discussed however it yields different uncertainty relations than the ones we consider and does not result in a single uncertainty relation in higher dimensions.

In Theorem 6.3.12 we apply Theorem 6.3.2 to the Riesz transform of a function and get a Heisenberg uncertainty relation for the Riesz transformed function. Furthermore we show that while the frequency localization is invariant under the Riesz transform the localization in space domain is not and we give a formula for the localization in space domain.

As a further application of Theorem 6.3.3 we give an affine uncertainty relation Theorem 6.3.13 that is based on the tensor approach to the affine group and is to our knowledge entirely new.

Chapter 7

The Gibbs phenomenon for wavelet frames

7.1 Preliminaries

We like to investigate the existence of Gibbs' phenomenon for wavelet frames which are based on some kind of multiresolution analysis, i.e., for which a scaling function exists that satisfies condition (7.1).

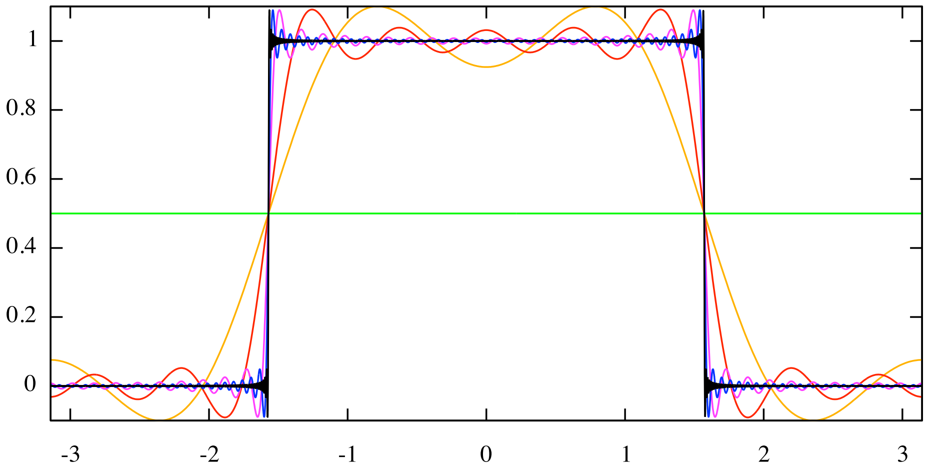


Figure 7.1: The Gibbs phenomenon demonstrated by partial sums S_n of the Fourier series for the characteristic function of the interval $[-\pi/2, \pi/2]$: S_1 =green, S_5 =orange, S_{10} =red, S_{40} =magenta, S_{100} =blue, S_{500} =black

The Gibbs phenomenon was first described for partial sums of Fourier series. (See [27] for further details.) It demonstrates the difficulties in approximating functions with jump discontinuities using partial Fourier series. The Gibbs phenomenon describes the appearance of overshoots and undershoots near jump discontinuities as shown in Figure 7.1.

Here the characteristic function of the interval $[-\pi/2, \pi/2]$ is approximated by the partial Fourier series

$$S_n(x) := \sum_{k=0}^{n} \int_{[-\pi/2,\pi/2]} e^{-2\pi ikt} dt e^{2\pi ikx}, \ \forall x \in [-\pi, \pi]$$

This phenomenon is not restricted to Fourier series but also occurs in other representations of functions containing jump discontinuities such as the representations using general orthogonal series expansions, general integral transforms, spline approximations, and continuous as well as discrete wavelet approximations. (See [27] for further details.)

In [44] the Gibbs phenomenon for orthonormal wavelet bases based on a MRA is studied. Our approach is to extend certain results of [44] from orthonormal wavelet bases to a certain class of wavelet frames.

The main contribution here is the observation that condition (7.1) is needed for the proof that the Gibbs phenomenon occurs. It is automatically satisfied by wavelet orthonormal bases.

In order to state a formal definition of the Gibbs phenomenon we will need some notation. Let $A \in \mathrm{GL}(n)$ and let

$$\{\psi_{j,k}\}_{j\in\mathbb{Z},k\in\mathbb{Z}^n} := \{A^j T_k \psi\}_{j\in\mathbb{Z},k\in\mathbb{Z}^n} \subset L^2(\mathbb{R}^n)$$

be a wavelet frame with dual frame

$$\{\psi^d_{j,k}\}_{j\in\mathbb{Z},k\in\mathbb{Z}^n} := \{A^j T_k \psi^d\}_{j\in\mathbb{Z},k\in\mathbb{Z}^n} \subset L^2(\mathbb{R}^n)$$

such that there exist scaling functions $\phi, \phi^d \in L^2(\mathbb{R}^n)$ which satisfy

$$\sum_{j\in\mathbb{N}} \sum_{k\in\mathbb{Z}^n} \langle f, A^{-j} T_k \psi\rangle A^{-j} T_k \psi^d = \sum_{k\in\mathbb{Z}^n} \langle f, T_k \phi\rangle T_k \phi^d. \tag{7.1}$$

Remark 7.1.1 (Sufficient conditions for (7.1))
LET $\phi \in L^2(\mathbb{R}^n)$ *be a refinable function with refinement mask* H_0 *such that* $\widehat{\phi}$ *is continuous at the origin and* $\lim_{|\xi|\to 0} \widehat{\phi}(\xi) = 1$. *Furthermore for let* $H_1 \in L^\infty(\mathbb{T})$ *be the wavelet mask for* $\psi \in L^2(\mathbb{R}^n)$, *where* ψ *is the mother wavelet of a tight wavelet frame with frame bound* $A = 1$.

Moreover, suppose that for all $q \in \mathfrak{D}_a^{-1}\mathbb{Z}^n/\mathbb{Z}^n$ *and for almost all* $\xi, \xi + q \in \mathbb{T}$ *with*

$$\sum_{k\in\mathbb{Z}^n} |\widehat{\phi}(\xi + k)|^2 > 0 \quad \wedge \quad \sum_{k\in\mathbb{Z}^n} |\widehat{\phi}(\xi + q + k)|^2 > 0$$

equation (5.9) *hold:*

$$H_0(\xi) H_0(\xi + q) + H_1(\xi) H_1(\xi + q) = \delta_{q,0}.$$

THEN *Theorem 5.5.14 states that condition* (7.1) *is satisfied, where* $\phi = \phi^d$ *and* $\psi = \psi^d$.

Proposition 7.1.2
LET $\phi, \phi^d, \psi, \psi^d$ *be such that condition* (7.1) *holds. Let*

$$q(x,t) := \sum_{k\in\mathbb{Z}^n} \phi(x-k)\overline{\phi^d(t-k)}$$

and $q_m(x,t) := |det(A)|^m q(A^m x, A^m t)$.

THEN

$$\begin{aligned} f_m(x) :&= \int_{\mathbb{R}^n} q_m(x,t) f(t) dt = \int_{\mathbb{R}^n} \sum_{k\in\mathbb{Z}^n} \phi(A^m x - k)\overline{\phi^d(A^m t - k)} f(t) dt \\ &= \sum_{k\in\mathbb{Z}^n} \langle f, A^m T_k \phi\rangle A^m T_k \phi^d(x) \overset{(7.1)}{=} \sum_{k\in\mathbb{Z}^n, j<m} \langle f, A^j T_k \psi\rangle A^j T_k \psi^d(x) \end{aligned}$$

and hence $f_m \overset{m\to\infty}{\longrightarrow} f$.

7.2 The Gibbs phenomenon

EXAMPLE 7.2.1 (Gibbs phenomenon in one dimension):
LET $f \in L^2(\mathbb{R})$ be continuous in a neighborhood of some $a \in \mathbb{R}$ except for a jump at a. Let

$$a_- := \lim_{t\to a, t<a} f(t) \quad \text{and} \quad a_+ := \lim_{t\to a, t>a} f(t).$$

Without loss of generality let $a_- < a_+$.
THEN the dual frames generated by $\psi, \psi^d \in L^2(\mathbb{R})$ are said to exhibit a **Gibbs phenomenon** at a if either $\exists\{a_l\}_{l\in\mathbb{N}} : a_l < a$ and $\lim_{m\to\infty} f_m(a_m) < a_-$ or $\exists\{a_l\}_{l\in\mathbb{N}} : a_l > a$ $\lim_{m\to\infty} f_m(a) > a_+$.

In the case that the dimension is $n > 1$ we will use the following definition of the Gibbs phenomenon.

Definition 7.2.1 (Gibbs phenomenon)
Let $f : \mathbb{R}^n \to \mathbb{R}$ have the property that there exists a $(n-1)$–dimensional hypersurface S satisfying $0 \in S$ and that f is continuous $\forall x \in \mathbb{R}^n \setminus S$. Let this hypersurface S be given by a function $g \in C^1(\mathbb{R}^n)$ with the property that $\nabla g(0) \neq 0$ via $g(x) = 0 \Leftrightarrow x \in S$. It follows, that $g(0) = 0$ and f is continuous $\forall x \in \mathbb{R}^n$ such that $g(x) \neq 0$.

There exists a radius $r > 0$ such that the surface S divides the ball $B_r(0)$ into two connected parts $B^+ := \{x \in B_r(0) : g(x) > 0\}$ and $B^- := \{x \in B_r(0) : g(x) \leq 0\}$.

f is called **piecewise continuous** *on $B_r(0)$ if there exist two continuous functions f^+, f^- on $B_r(0)$ such that $f(x) = f^\pm(x)$, $\forall x \in B^\pm$.*

f is said to have a **jump discontinuity at** 0 *if $f^+(0) \neq f^-(0)$.*

LET *$f : \mathbb{R}^n \to \mathbb{R}$ be a be a bounded, measurable function which is piecewise continuous and exhibits a jump discontinuity at 0.*

Let $\nabla g(x) := (\frac{\partial}{\partial x_1} g(x), \dots, \frac{\partial}{\partial x_n} g(x))$ and

$$\gamma := \frac{\nabla g(0)}{|\nabla g(0)|}. \tag{7.2}$$

Let μ be a unit vector such that $\langle \gamma, \mu \rangle \neq 0$.

THEN *the wavelet approximation f_m is said to show a* **Gibbs phenomenon at** 0 **in the direction of** μ, *if there exists a sequence $\{a_m\}_{m\in\mathbb{N}} \subset \mathbb{R}^+$ such that $a_m \to 0$ as $m \to \infty$ and*

$$\begin{array}{lll} \textit{either} & L > \lim_{m\to\infty} f(a_m\mu), & \textit{when } \lim_{m\to\infty} f(a_m\mu) > \lim_{m\to\infty} f(-a_m\mu) \\ \textit{or} & L < \lim_{m\to\infty} f(a_m\mu), & \textit{when } \lim_{m\to\infty} f(a_m\mu) < \lim_{m\to\infty} f(-a_m\mu), \end{array}$$

where $L = lim_{m\to\infty} f_m(a_m\mu)$.

Definition 7.2.2 (l-regular functions)
*A function $f \in C^l(\mathbb{R}^n)$ is called l-**regular** in the sense of Mallat, or simply l-regular, if*

$$\forall \boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}_0^n,\ |\boldsymbol{\alpha}| \leq l,\ \forall k \in \mathbb{N}\, \exists C_k > 0 : |\partial^{\boldsymbol{\alpha}} f(x)| \leq C_k(1 + |x|)^{-k},\ \forall x \in \mathbb{R}^n.$$

Henceforth we will assume that ϕ, ϕ^d are l-regular for some $l \in \mathbb{N}_0$.
It follows that

$$\forall |\boldsymbol{\alpha}|, |\boldsymbol{\beta}| \leq l,\ \forall k \in \mathbb{N}\, \exists C_k > 0 : |\partial_x^{\boldsymbol{\alpha}} \partial_t^{\boldsymbol{\beta}} q(x,t)| \leq C_k(1 + |x - t|)^{-k},\ \forall x, t \in \mathbb{R}^n. \tag{7.3}$$

To show that the formulation of the Gibbs phenomenon is meaningful we show that the Gibbs phenomenon does not occur at points, where the function f is continuous. This theorem extends a similar theorem for wavelet orthonormal bases given in [44].

Theorem 7.2.3
LET $f:\mathbb{R}^n \to \mathbb{R}$ *be bounded and measurable. Let* $K \subset \mathbb{R}^n$ *be compact and let* f *be continuous for every* $x \in K$.
THEN $f_m(x) \overset{m\to\infty}{\longrightarrow} f(x)$, *uniformly* $\forall x \in K$.

Proof. By the regularity of q it is clear, that $f_m(x)$ converges uniformly, whence $\lim_{m\to\infty} f_m$ is a continuous function. Indeed let $C := \sup_{x\in\mathbb{R}^n} \int_{\mathbb{R}^n} |q(x,t)|dt$. Since K is compact and f is continuous on K,

$$\forall \epsilon > 0 \exists \delta > 0 : |f(t) - f(x)| < \frac{\epsilon}{2C}, \forall x \in K, t \in \mathbb{R}^n : |x - t| < \delta.$$

For $x \in K$

$$\begin{aligned}
|f_m(x) - f(x)| &= \Big| \int_{\mathbb{R}^n} q_m(x,t)(f(t) - f(x))dt \Big| \\
&\le \int_{|t-x|<\delta} |q_m(x,t)||f(t) - f(x)|dt + \int_{|t-x|\ge\delta} |q_m(x,t)||f(t) - f(x)|dt \\
&= I + J
\end{aligned}$$

Now we estimate I as

$$\begin{aligned}
I &= \int_{|t-x|<\delta} |q_m(x,t)||f(t) - f(x)|dt \le \frac{\epsilon}{2C} \int_{|t-x|<\delta} |\det(A^m)||q(A^m x, A^m t)|dt \\
&= \frac{\epsilon}{2C} \int_{|A^{-m}t-x|<\delta} |q(A^m x, t)|dt \\
&\le \frac{\epsilon}{2}.
\end{aligned}$$

Choosing $k > n$ in (7.3), we continue by estimating J by

$$\begin{aligned}
J &= \int_{|t-x|\ge\delta} |q_m(x,t)||f(t) - f(x)|dt \le \int_{|t-x|\ge\delta} |\det(A^m)||q(A^m x, A^m t)|2\|f\|_\infty dt \\
&\le |\det(A^m)| \int_{|t-x|\ge\delta} \frac{C_k}{(1 + |A^m(t-x)|)^k} 2\|f\|_\infty dt \\
&= \int_{|A^{-m}t|\ge\delta} \frac{C_k}{(1+|t|)^k} 2\|f\|_\infty dt \overset{m\to 0}{\longrightarrow} 0.
\end{aligned}$$

It follows that for any $\epsilon > 0$ there exists some $m_0 \in \mathbb{N}$ such that $|f_m(x) - f(x)| < \epsilon$, $\forall m \ge m_0$, $x \in K$. □

7.3 Sufficient conditions for the occurrence of the Gibbs phenomenon

The following theorem states sufficient conditions for a Gibbs phenomenon to occur for a certain very simple model function.

Theorem 7.3.1 (Sufficient conditions)
LET $\phi, \phi^d : \mathbb{R}^n \to \mathbb{R}$ *be refinable functions such that* (7.1) *is fullfilled. Let* μ, *and* γ, *be eigenvectors of* A, *and* A^*, *respectively, with the same positive eigenvalue, which fullfill* $|\mu| = |\gamma| = 1$ *and* $\langle \gamma, \mu \rangle > 0$. *Finally let* $f(x) = H(\langle \gamma, x \rangle)$, *where* $H(x) := \begin{cases} 1 & x \geq 0 \\ 0 & x < 0 \end{cases}$ *is the Heaviside function.*

THEN *the wavelet approximation* f_m *exhibits a Gibbs phenomenon at* 0, *iff*

$$\int_{\langle \gamma, t \rangle > 0} q(a\mu, t)\, dt > 1 \text{ for some } a > 0. \tag{7.4}$$

Proof. The proof is the same as in the orthonormal bases case found in [44].

$$f_m(x) = \int_{\langle \gamma, t \rangle > 0} |\det(A^m)| q(A^m x, A^m t)\, dt = \int_{\langle \gamma, A^{-m} v \rangle > 0} q(A^m x, v)\, dv$$

We know that $\exists \lambda > 0 :\ A^* \gamma = \lambda \gamma$ and $A\mu = \lambda \mu$. Since

$$\langle \gamma, v \rangle = \left\langle (A^*)^m \gamma, A^{-m} v \right\rangle = \left\langle \lambda^m \gamma, A^{-m} v \right\rangle,$$

we have that for every $b \in \mathbb{R}$

$$f_m(b\mu) = \int_{\langle \gamma, v \rangle > 0} q(b\lambda^m \mu, v)\, dv. \tag{7.5}$$

Let us assume that (7.4) holds. Set $b = a_m = a\lambda^{-m}$ to obtain

$$f_m(a_m \mu) = \int_{\langle \gamma, v \rangle > 0} q(a\mu, v)\, dv > 1.$$

Hence $f_m(a_m \mu)$ is a constant sequence whith value greater than 1. Since

$$1 = \lim_{t \to 0,\, t > 0} f(t) > \lim_{t \to 0,\, t < 0} f(t) = 0$$

we have shown a Gibbs phenomenon in direction μ to exist.

Conversely assume, that

$$\exists \{a_m\}_{m \in \mathbb{N}} : 0 > a_m \overset{m \to \infty}{\longrightarrow} 0, \text{ such that } \lim_{m \to \infty} f_m(a_m \mu) > 1.$$

Then $\exists m \in \mathbb{N}$, $f_m(a_m \mu) > 1$. Setting $b = a^m \lambda^{-m}$ in (7.5) we get (7.4). □

Lemma 7.3.2
LET $f : \mathbb{R}^n \to \mathbb{R}$ *be bounded and measurable, and let* A *be a dilation matrix all of whose eigenvalues have the same absolute value. Consider the cone*

$$C := \left\{ x \in \mathbb{R}^n : |\langle x, \mu \rangle| \geq |x| \cos(\theta) \right\},$$

where $\theta \in]0, \frac{\pi}{2}]$ *and* $\mu \in \mathbb{R}^n$, $|\mu| = 1$. *Furthermore, let* $f(x) = 0$, $\forall x \in C$, *and* $\lim_{b \to 0, b > 0} f(by) = 0$, *for almost all* $y \in \mathbb{R}^n$.

THEN $f_m(a\mu) \overset{m \to \infty}{\longrightarrow} 0$ *uniformly for* $a \in \mathbb{R}$.

Proof.

$$f_m(x) = \int_{\mathbb{R}^n} q_m(x,t) f(t)\,dt = \int_{\mathbb{R}^n \setminus C} q_m(x,t) f(t)\,dt.$$

Let $x = a\mu$ with $a \in \mathbb{R}$. Let $t \in \mathbb{R}^n \setminus C$ and $k > n$.
Then

$$\begin{aligned}|x-t| &= \sqrt{a^2 + |t|^2 - 2a\langle\mu, t\rangle} \geq \sqrt{a^2 + |t|^2 - 2|a|\,|\langle\mu, t\rangle|} \\ &\geq \sqrt{a^2 + t^2 - 2a|t|\cos(\theta)} \geq \sqrt{a^2 + t^2 - a^2 - t^2\cos^2(\theta)} \\ &= |t|\sin(\theta)\end{aligned}$$

and consequently

$$\begin{aligned}|q_m(x,t)| &\overset{(7.3)}{\leq} C_k \big|\det(A^m)\big|\big(1 + |A^m(x-t)|\big)^{-k} \leq C_k \big|\det(A^m)\big|\big(1 + \big|\det(A^m)\big||x-t|\big)^{-k} \\ &\leq C_k |\det(A^m)|\big(1 + \big|\det(A^m)\big|\,|t|\sin(\theta)\big)^{-k}.\end{aligned}$$

Applying this to f_m we get

$$\begin{aligned}|f_m(a\mu)| \leq& C_k \int_{\mathbb{R}^n} \big|\det(A^m)\big|\big(1 + \big|\det(A^m)\big|\,|t|\sin(\theta)\big)^{-k} f(t)\,dt \\ =& C_k \int_{\mathbb{R}^n} (1 + |t|\sin(\theta))^{-k} f\big(\big|\det(A^m)\big|^{-1} t\big)\,dt.\end{aligned}$$

By the Lebesgue dominated convergence theorem the last expression converges uniformly to zero as $m \to \infty$. □

Theorem 7.3.3
LET $f : \mathbb{R}^n \to \mathbb{R}$ *be a bounded measurable function that is piecewise continuous at* 0*. Let* γ *as in* (7.2) *and let* $\mu \in \mathbb{R}^n : |\mu| = 1$, $\langle\mu,\gamma\rangle > 0$.

THEN $\lim_{m\to\infty} f_m$ *exhibits a Gibbs phenomenon at* 0 *iff* (7.4) *holds.*

Proof. Let $r, B^{\pm}$ and $f^{\pm}$ be as in Definition 7.2.1. Replacing f by a linear combination $\alpha f(x) + \beta$ if necessary, we may assume that $f^+(0) = 1$ and $f^-(0) = 0$. Let $\theta \in]0, \frac{\pi}{2}[: \theta < \frac{\pi}{2} - \psi$, where ψ is the angle between γ and μ. Let

$$C^{\pm} := \big\{x \in \mathbb{R}^n : \big|\langle x,\mu\rangle\big| \geq \cos(\theta)|x| \text{ and } \pm\langle x,\mu\rangle > 0\big\}.$$

$C^{\pm}$ lies on opposite sides of the tangent plane $\gamma^{\perp} = \big\{x \in \mathbb{R}^n : \langle x,\gamma\rangle = 0\big\}$. Then $C^{\pm} \cap B_r(0) \subset B^{\pm}$.
Finally let $F : \mathbb{R}^n \to \mathbb{R},\ x \mapsto \begin{cases} 1, & \forall x \in B^+ \cup C^+; \\ 0, & \text{otherwise.}\end{cases}$

Then there exists an $\epsilon > 0$ such that $f - F$ is continuous for all $a\mu$ such that $a < \epsilon$. Therefore, by Theorem 7.2.3 and by the linearity of the wavelet transform, f_m shows a Gibbs phenomenon in the direction of μ iff F_m does. Now consider $G(x) := F(x) - H(\gamma \cdot x)$. Then G satisfies all hypotheses for Lemma 7.3.2, hence $G_m(a\mu) \overset{m\to 0}{\longrightarrow} 0$ uniformly for $a \in \mathbb{R}$. Consequently f_m exhibits a Gibbs phenomenon iff $H_m(\gamma \cdot\ \cdot\ )$ does. Now Theorem 7.3.1 gives the result.

□

Chapter 8

Higher Riesz transforms and represenations of the rotation group

A wavelet system consists of a set of functions derived from a mother wavelet by theapplication of a set of operators which are a unitary representation of a certain group - in the one dimensional case this is the affine group on $L^2(\mathbb{R})$. These sets are then sampled to give a basis or a frame for $L^2(\mathbb{R}^n)$. Riesz transforms are very well suited for an implementation via wavelets. In this chapter we show that this is due to the fact that the steering corresponds to a unitary representation of the rotation group on the set of Riesz transformed functions. For $n > 1$ there are an infinite number of such unitary representations of $\mathrm{SO}(n)$ based on the spherical harmonics which yield steerable operators - the higher Riesz transforms. We give explicit constructions of higher Riesz transforms for the interesting cases $n = 2, 3$.

New contributions in this chapter are first of all the definition of higher Riesz transforms as the vector of partial higher Riesz transforms Definition 8.2.4 corresponding to a basis of spherical harmonics $\mathscr{H}_k$. These higher Riesz transforms are the smallest possible sets of partial higher Riesz transforms which are invertible and steerable. Combining these minimal higher Riesz transforms in Corollary 8.2.18 we derive higher Riesz transforms which are geometrically richer at the cost of higher redundancy. In Remark 8.2.19 we learn how to construct Higher Riesz transforms which contain the directional information of higher derivatives. This approach is similar to the steerable pyramids of Freeman and Adelson in [18] – indeed in Remark 8.2.22 we show that these steerable pyramids have a close connection to higher Riesz transforms.

To proof that the higher Riesz transforms map frames onto frames in Theorem 8.2.7 we first show that the hypercomplex higher Riesz transforms defined in Definition 8.2.4 are unitary operators. Using our result Theorem 5.2.13, we conclude in Theorem 8.2.8 that the higher Riesz transforms map Clifford frames onto Clifford frames.

A key property for image analysis of the Riesz transform is the decomposition of an image into amplitude, phase and phase direction. In Definition 8.2.9 we state the concept of phase direction for higher Riesz transforms. Based on the concept of hypercomplex higher Riesz transforms we define a higher monogenic signal in Definition 8.2.11 which yields an amplitude-phase decomposition for higher Riesz transforms.

Theorem 8.2.15 yields the connection of the higher monogenic signal to certain generalized Cauchy Riemann equations derived from higher Dirac operators Definition 8.2.13 which are the square-root of a power of the Laplace operator.

8.1 Definitions and basic theorems

We will start with some basic definitions and theorems which can be found in similar form in [33, 45, 14] and [57].

8.1.1 Definition of spherical harmonics

Definition 8.1.1 (Linear spaces of homogenous polynomials)
A function f on a vector space V is called **homogenous of degree** $k \in \mathbb{N}$ *iff*

$$\forall \epsilon > 0 \; f(\epsilon x) = \epsilon^k f(x), \; \forall x \in V.$$

Let $V = \mathbb{R}^n$. The space of **homogenous polynomials** *of order k is denoted by*

$$\mathfrak{P}_k(\mathbb{R}^n) := \Big\{ p(x) = \sum_{|\alpha|=k} a_\alpha x^\alpha, p \text{ is homogenous of degree } k \Big\},$$

where $\alpha \in \mathbb{N}_0^n$ is a multiindex and n is the dimension of the vector space V.
Let $f \in C^2(V)$. The **Laplace differential** *operator is given by*

$$\triangle f(x) = \sum_{l=1}^{n} \frac{d^2}{dx_k^2} f(x), \; \forall x \in V.$$

The **solid spherical harmonics** *are the subspace of the homogenous polynomials of degree k that lies in the kernel of $\triangle$:*

$$\mathfrak{H}_k(\mathbb{R}^n) = \{ p \in \mathfrak{P}_k : \triangle p = 0 \}.$$

The **surface spherical harmonics** *are the restriction of the solid spherical harmonics to the unit sphere:*

$$\mathcal{H}_k := \{ p|_{S^{n-1}}, \; p \in \mathfrak{H}_k \}.$$

Remark 8.1.2
For the rest of this chapter $k \in \mathbb{N}_0$ will denote the order of the polynomial spaces considered.
Of course $\mathcal{H}_k$ is a linear subspace of $L^2(S^{n-1})$.
A scalar product on $L^2(S^{n-1})$ is given by

$$\langle P, Q \rangle := \int_{S^{n-1}} P(x) \overline{Q(x)} \omega_n^{-1} d\sigma(x), \; \forall P, Q \in L^2(S^{n-1}),$$

where $\omega_n = \frac{2\pi^{n/2}}{\Gamma(n/2)}$ *is the surface area of the unit sphere.*

8.1.2 Properties of spherical harmonics

Theorem 8.1.3 (Orthogonality of spherical harmonics)
The finite dimensional spaces $\mathcal{H}_k$, $k \in \mathbb{N}_0$ are mutually orthogonal with respect to the scalar product on $L^2(S^{n-1})$.

Proof. This is 3.1.1 in [45]. □

The following theorem gives a relation between homogenous polynomials and spherical harmonics.

Theorem 8.1.4 (Dimension of spherical harmonics)
LET $p \in \mathfrak{P}_k$.
THEN *there exist* $p_1 \in \mathscr{H}_k$, $p_2 \in \mathfrak{P}_{k-2}$ *such that* $p(x) = p_1(x) + |x|^2 p_2(x)$, $\forall x \in \mathbb{R}^n$. *The dimension of* $\mathfrak{P}_k$ *is*

$$\dim(\mathfrak{P}_k) = \binom{n+k-1}{k},$$

whence the dimension of the spaces of spherical harmonics is

$$d_{n,k} := \dim(\mathfrak{H}_k(\mathbb{R}^n)) = \dim(\mathscr{H}_k(\mathbb{R}^n)) = \binom{n+k-1}{k} - \binom{n+k-3}{k-2}.$$

Proof. This is 3.1.2 in [45]. □

Theorem 8.1.5 (Completeness of spherical harmonics)
The collection of all finite linear combinations of elements of $\mathscr{H}_k(\mathbb{R}^n)$ *is*

(i) dense in $\left(C(S^{n-1}), \|\cdot\|_\infty\right)$;

(ii) dense in $L^2(S^{n-1})$.

Proof. This is IV.2.3 in [46]. □

Definition 8.1.6 (Some orthogonal polynomials)
The **Gegenbauer polynomials** *also called* **ultraspherical polynomials** C_k^ν *of degree* k *and order* $\nu \neq 0$ *are generated by the function*

$$(1 - 2xt + t^2)^{-\nu} = \sum_{k=0}^{\infty} C_k^\nu(x) t^k, \ \forall x, t \in \mathbb{R}.$$

The **Legendre polynomial** *of degree* k *is defined by*

$$P_k(x) = \frac{2^{-k}}{k!} \frac{d^k}{dx^k} (x^2 - 1)^k, \ \forall x \in \mathbb{R}.$$

The **Tchebichef polynomial** *of degree* k *is generated by*

$$-1/2 \log(1 - 2xt + t^2) = \sum_{k \in \mathbb{N}_0} (k+1)^{-1} T_{k+1}(x) t^{k+1}, \ \forall x, t \in \mathbb{R}.$$

Theorem 8.1.7 (Explicit formulas for orthogonal polynomials)

(i) LET $2\nu \in \mathbb{N}$ *and let* $l = \lfloor \nu \rfloor \in \mathbb{N}_0$.

THEN

$$C_k^\nu(x) = C_k^{l+1/2}(x) = \frac{2^l l!}{(2l)!} \frac{d^l}{dx^l} P_{n+l}(x) = \frac{2^{-k} l!}{(2l)!(k+l)!} \frac{d^{k+2l}}{dx^{+2l}} (x^2 - 1)^{k+l}, \ \forall x \in \mathbb{R},$$

respectively,

$$C_k^\nu(x) = C_k^{l+1}(x) = \frac{2^{-l}}{l!(k+l+1)} \frac{d^{l+1}}{dx^{l+1}} T_{k+l+1}(x), \ \forall x \in \mathbb{R}.$$

(ii) LET $l \in \mathbb{N}_0$.

THEN

$$C_k^{l/2}(1) = \frac{(k+l-1)!}{k!(l-1)!} = \binom{k+l-1}{k}.$$

(iii) The Tchebichef polynomials are given as

$$T_k(x) = \frac{1}{2}\Big(\big(x + i(1-x^2)^{\frac{1}{2}}\big)^k + \big(x - i(1-x^2)^{\frac{1}{2}}\big)^k\Big) = \cos\big(k\cos^{-1}(x)\big), \ \forall x \in \mathbb{R}.$$

Theorem 8.1.8 (Addition theorem)
LET $\{y_k^l\}_{l=1}^{d_{n,k}} \subset H_k$ *be a real orthonormal basis of* $\mathscr{H}_k$ *with respect to the scalar product on* $L^2(S^{n-1})$*.*
THEN *for any fixed* $\eta \in S^{n-1}$

$$\frac{C_k^{\frac{n-2}{2}}(\langle \xi, \eta\rangle)}{C_k^{\frac{n-2}{2}}(1)} = \frac{1}{d_{n,k}} \sum_{l=1}^{d_{n,k}} y_k^l(\xi) y_k^l(\eta), \ \forall \xi \in S^{n-1}.$$

Especially

$$\sum_{l=1}^{d_{n,k}} |y_k^l(\xi)|^2 = d_{n,k}. \tag{8.1}$$

Proof. See [14], Chapter 11.4. □

Lemma 8.1.9 (Representation of the rotation group)
LET $\big\{y_k^l\big\}_{l=1}^{d_{n,k}} \subset \mathscr{H}_k$ *be an orthonormal basis of* $\mathscr{H}_k$ *and let* $\rho \in SO(n)$*.*
THEN $\exists \mathfrak{D}_{\rho,k} = (D_{l,r})_{l,r=1}^{d_{n,k}} \in SO(d_{n,k})$ *such that*

$$y_k^l(\rho\xi) = \sum_{r=1}^{d_{n,k}} \mathfrak{D}_{l,r}\, y_k^r(\xi).$$

$\mathfrak{D}_{\rho,k}$ *is a representation of the rotation group.*

Proof. See [14], Chapter 11 Lemma 5. □

8.1.3 The spaces $H_k(\mathbb{R}^n)$

Definition 8.1.10 (The spaces $H_k(\mathbb{R}^n)$)
$H_k(\mathbb{R}^n) := \big\{f \in L^2(\mathbb{R}^n) : \ f(x) = g(|x|)h_k(x), \ \textit{where}\ h_k \in \mathfrak{H}_k(\mathbb{R}^n), \ g : \mathbb{R}^+ \to \mathbb{R}\big\}.$

Theorem 8.1.11 (Completeness of the spaces H_k)
It holds that $L^2(\mathbb{R}^n) = \bigoplus_{k\in\mathbb{N}_0} H_k(\mathbb{R}^n)$ *in the following sense:*

(i) Let $\{\mathfrak{h}_1^k,\dots,\mathfrak{h}_{d_{n,k}}^k\}$ *be an orthonormal basis of* $\mathfrak{H}_k$.
Then any $f\in H_k(\mathbb{R}^n)$ *can be written as*

$$f(x) = \sum_{j=1}^{d_{n,k}} f_j(|x|)\mathfrak{h}_j^k(x).$$

It follows that

$$\int_{\mathbb{R}^n} |f(x)|^2 dx = \sum_{j=1}^{d_{n,k}} \int_{\mathbb{R}_0^+} |f_j(r)|^2 r^{2k+n-1} dr.$$

(ii) Each subspace $H_k(\mathbb{R}^n)$ *is closed.*

(iii) The $H_k(\mathbb{R}^n)$ *are mutually orthogonal.*

(iv) Every element of $L^2(\mathbb{R}^n)$ *is a limit of finite linear combinations of elements belonging to the spaces* $H_k(\mathbb{R}^n)$.

Proof. This is 2.1.8 in [46]. □

Corollary 8.1.12 (Invariance under the Fourier transform)
The spaces H_k *are invariant under the Fourier transform, i.e.,* $\mathscr{F}: H_k(\mathbb{R}^n)\to H_k(\mathbb{R}^n)$.
LET $p\in\mathfrak{H}_k$ *and* $f:\mathbb{R}\to\mathbb{R}$ *such that* $pf(|\cdot|)\in L^2(\mathbb{R}^n)$.
THEN *there exists a function* $g:\mathbb{R}\to\mathbb{R}$ *such that*

$$\mathscr{F}(pf)(\xi) = p(\xi)g(|\xi|),\ \text{f.a.a. } \xi\in\mathbb{R}^n.$$

Proof. This is a corollary of Theorem 3.4 in [45]. □

8.2 Higher Riesz transforms and irreducible representations of the rotation group

Remark 8.2.1 (An algebra of Riesz transforms)
Since the partial Riesz transforms R_α, $\alpha\in\{1,\dots,n\}$, *map* $L^p(\mathbb{R}^n)$ *to* $L^p(\mathbb{R}^n)$ *for* $1<p<\infty$, *they span an algebra* $\mathfrak{A}$ *of linear bounded operators, invariant under translation and dilation. Let* $\boldsymbol{\alpha},\boldsymbol{\beta}\in\mathbb{N}_0^n$, $a\in\mathbb{R}$. *The algebra is generated by*

$$\cdot : \mathfrak{A}\times\mathfrak{A}\to\mathfrak{A}, \qquad \mathscr{F}(R^{\boldsymbol{\alpha}}\cdot R^{\boldsymbol{\beta}}f)(\xi) = \frac{i^{|\boldsymbol{\alpha}|}\xi^{\boldsymbol{\alpha}}\xi^{\boldsymbol{\beta}}}{|\xi|^{|\boldsymbol{\alpha}|+|\boldsymbol{\beta}|}}\widehat{f}(\xi),$$

$$+ : \mathfrak{A}\times\mathfrak{A}\to\mathfrak{A}, \qquad \mathscr{F}\big((R^{\boldsymbol{\alpha}}+R^{\boldsymbol{\beta}})f\big)(\xi) = \Big(\frac{i^{|\boldsymbol{\alpha}|}\xi^{\boldsymbol{\alpha}}}{|\xi|^{|\boldsymbol{\alpha}|}} + \frac{i^{|\boldsymbol{\beta}|}\xi^{\boldsymbol{\beta}}}{|\xi|^{|\boldsymbol{\beta}|}}\Big)\widehat{f}(\xi),$$

and multiplication with a scalar is given by

$$\cdot : \mathbb{R}\times\mathfrak{A}\to\mathfrak{A}, \qquad \widehat{aR^{\boldsymbol{\alpha}}f}(\xi) = a\frac{i^{|\boldsymbol{\alpha}|}\xi^{\boldsymbol{\alpha}}}{|\xi^{|\boldsymbol{\alpha}|}|}f(\xi),\ \forall f\in L^p(\mathbb{R}^n),\ \text{f.a.a. } \xi\in\mathbb{R}^n.$$

It is evident that the elements of the algebra $\mathfrak{A}$ are invariant with respect to dilation and translation. Examples of elements of this algebra are the operators defined in Theorem 8.2.2. In this section we find rotation invariant linear subspaces of $\mathfrak{A}$ on which there exists a representation of the rotation group.

Theorem 8.2.2 (Fourier multipliers of singular integral operators)
LET $P \in \mathfrak{H}_k(\mathbb{R}^n)$, $k \in \mathbb{N}$, $1 < p < \infty$. *Furthermore, let the singular integral operator*

$$\mathscr{R}_P : L^p(\mathbb{R}^n) \to L^p(\mathbb{R}^n)$$

be given by

$$\mathscr{R}_P f(x) := \lim_{\epsilon \to 0} \int_{|y| \geq 0} \gamma_k \frac{P(y)}{|y|^{k+n}} f(x-y)\,dy, \ \forall f \in L^p(\mathbb{R}^n), \text{ f.a.a. } x \in \mathbb{R}^n,$$

where $\gamma_k := \frac{\Gamma\left(\frac{k+n}{2}\right)}{\pi^{n/2}\Gamma\left(\frac{k}{2}\right)}$.
THEN *the Fourier multiplier corresponding to* $\mathscr{R}_P$ *is*

$$m_p(\xi) := i^k \frac{P(\xi)}{|\xi|^k}, \ \forall \xi \in \mathbb{R}^n.$$

That is $\mathscr{F}\left(\mathscr{R}_P f\right) = m_p \widehat{f}$.

Proof. This is theorem 3.5 in [45]. □

Remark 8.2.3 (Riesz transforms depend on surface spherical harmonics)
Every solid spherical harmonic $p \in \mathfrak{H}_k$ *may be written at a point* $x \in \mathbb{R}^n$ *as* $p(x) = |x|^k p\left(\frac{x}{|x|}\right)$. *The Fourier multiplier of the operator* $\mathscr{R}_p$ *from Theorem 8.2.2 is* $\frac{p(x)}{|x|^k} = p\left(\frac{x}{|x|}\right)$ *– the restriction to the unit sphere of the solid spherical harmonic* p *and as such it is a surface spherical harmonic. Hence the operator* $\mathscr{R}_p$ *from Theorem 8.2.2 depends only on the surface spherical harmonic* $p|_{S^{n-1}}$.

For the rest of the section we will assume that $\{\mathfrak{h}_1^k, \dots, \mathfrak{h}_{d_{n,k}}^k\}$ *is an ONB for* $\mathfrak{H}_k(\mathbb{R}^n)$ *in the sense that* $\langle \mathfrak{h}_l^k|_{S^{n-1}}, \mathfrak{h}_r^k|_{S^{n-1}} \rangle_{L^2(S^{n-1})} = \delta_{r,k}$, $\forall l, r = 1, \dots, d_{n,k}$. *It follows from Theorem 8.1.8 that*

$$\sum_{l=1}^{d_{n,k}} \left(\mathfrak{h}_l^k(x)\right)^2 = \sum_{l=1}^{d_{n,k}} \left(\mathfrak{h}_l^k\left(\frac{x}{|x|}\right)|x|^k\right)^2 = d_{n,k}|x|^{2k}, \ \forall x \in \mathbb{R}^n.$$

Definition 8.2.4 (Higher Riesz transform)
LET $P \in \mathfrak{H}_k(\mathbb{R}^n)$, $k \in \mathbb{N}$.

THEN *the singular integral operator* $\mathscr{R}_P$ *defined in Theorem 8.2.2 is called* **partial higher Riesz-transform**. k *is called the* **order** *of the higher Riesz transform.*

$\mathscr{R} = \mathscr{R}_k = d_{n,k}^{-1/2}\left(\mathscr{R}_{\mathfrak{h}_{l,k}}\right)_{l=1}^{d_{n,k}}$ *is called the* **higher Riesz transform** *of order* k. *The* **hypercomplex higher Riesz transform** *is the corresponding operator on the Clifford-Hilbert module* $L^2(\mathbb{R}^n)_{d_{n,k}}$ *given by*

$$\mathscr{R}_k : L^2(\mathbb{R}^n)_{d_{n,k}} \to L^2(\mathbb{R}^n)_{d_{n,k}}, \ f \mapsto \mathscr{R}_k f = d_{n,k}^{-1/2} \sum_{l=1}^{d_{n,k}} e_l \mathscr{R}_{\mathfrak{h}_l^k} f.$$

For notational convenience we will sometimes skip the index of $\mathscr{R}_k$ *and simply write* $\mathscr{R}$ *instead.*

Theorem 8.2.5
LET $\mathfrak{A}$ *be the algebra of operators on* $L^2(\mathbb{R}^n)$ *algebraically generated by the partial Riesz transforms* $R_1,\ldots,R_n$.
THEN

(i) every partial higher Riesz transform belongs to $\mathfrak{A}$*;*

(ii) the closure of $\mathfrak{A}$ *in the strong operator topology is identical with the algebra of bounded transformations on* $L^2(\mathbb{R}^n)$ *which commute with translations and dilations.*

Proof. This is 3.4.5 in [45]. □

8.2.1 Higher Riesz transforms and representations of the rotation group

Theorem 8.2.6 (Spherical harmonics and representations of the rotation group)
LET V *be a finite dimensional Hilbert space and let* $\rho \to \mathcal{T}_\rho$ *be a continuous homomorphism from* $SO(n)$ *to the group of unitary transformations on* V*. Recall from Definition 1.3.28 on page 14 that the couple* $(\mathcal{T}_\rho, V)$ *is called a unitary representation of* $SO(n)$*.*

A unitary representation is irreducible if there is no non-trivial invariant subspace of V *under* $\mathcal{T}_\rho$, $\rho \in SO(n)$*.*

Two representations $(\mathcal{T}_\rho, V_T)$ *and* $(\mathcal{S}_\rho, V_S)$ *are equivalent, iff there is a unitary correspondence* $U: V_T \leftrightarrow V_S$*, such that* $U^{-1}\mathcal{S}_\rho U = \mathcal{T}_\rho$*.*

(i) Let $V = \mathcal{H}_k$*. Then the following representation is irreducible:*

$$(\mathcal{T}_\rho P(x)) = P(\rho^{-1}x),\ \rho \in SO(n),\ P \in \mathcal{H}_k.$$

(ii) An irreducible representation $(\mathcal{S}_\rho, V)$ *of* $SO(n)$ *is equivalent to the one above iff*

$$\exists v \in V:\ \mathcal{S}_\rho(v) = v,\ \forall \rho \in SO(n-1).$$

(iii) Let $T: L^2(\mathbb{R}^n) \to L^2(\mathbb{R}^n, V)$ *be bounded and linear.*

(a) If T *commutes with translations and dilations and if*

$$\rho T \rho^{-1}(f) = \mathcal{T}_\rho T f,$$

then either $T = 0$ *or* $(\mathcal{T}_\rho, V)$ *is equivalent to the representation given in (i).*

(b) If $(\mathcal{T}_\rho, V)$ *arises from spherical harmonics of degree* $k \geq 1$ *then* T *is determined up to a constant multiple. Let* $\{\beta_1,\ldots,\beta_N\}$ *be a basis for the linear functionals on* V*, then each* $\beta_j(Tf)$ *is a partial higher Riesz transform of degree* k*.*

(iv) Let $\{\mathfrak{h}_l^k\}_{l=1}^{d_{n,k}} \subset \mathfrak{H}_k$ *be a orthonormal basis of* $\mathfrak{H}_k$*.* $\mathfrak{H}_k$ *is isomorphic to* $\mathbb{R}^{d_{n,k}}$ *via* $\mathfrak{h}_l^k \mapsto e_l$ *and* $\mathcal{T}_\rho$ *is isomorphic to a matrix* $\mathfrak{D}_\rho \in SO(d_{n,k})$ *- the Wigner* $\mathfrak{D}$*-matrix.*

Hence the Wigner $\mathfrak{D}$*-matrices yield a representation of the rotation group* $\left(\mathfrak{D}_\rho, \mathbb{R}^{d_{n,k}}\right)$*.*

Proof. (i) and (ii) are III.4.7 in [45]. (iii) is III.4.8 in [45].

□

EXAMPLE 8.2.1 (Higher Riesz transforms):
The spherical harmonics of degree k yield a number of spaces that have dimension $d_{n,k}$. In these spaces exist unitary representations of the rotation group given by unitary automorphisms that are equivalent to the Wigner $\mathfrak{D}$-matrices:

1. Let $U : \mathfrak{H}_k(\mathbb{R}^n) \to \mathbb{R}^{d_{n,k}},\ p = \sum_{\alpha=1}^{d_{n,k}} a_\alpha \mathfrak{h}_\alpha^k \mapsto (a_1, \ldots, a_{d_{n,k}})$.

   Then a irreducible representation $(\mathfrak{D}_\rho, \mathbb{R}^{d_{n,k}})$ of the group $SO(n)$ is given by
$$\mathscr{S}_\rho = U^{-1} \mathscr{T}_\rho U,$$
   where $(\mathscr{T}_\rho, \mathfrak{H}_k)$ is the representation given in 8.2.6(i). The matrices $\mathfrak{D}_\rho \in \mathrm{SO}(d_{n,k})$ were introduced in [57] by Wigner in and are called **Wigner-$\mathfrak{D}$ matrices**.

   This representation induces a representation $\big(\mathscr{S}_\rho, V\big)$:

2. Let $\mathscr{R} : L^2(\mathbb{R}^n) \to L^2(\mathbb{R}^n, \mathbb{R}^{d_{n,k}})$ be the higher Riesz transform given by
$$\mathscr{R}f = \{\mathscr{R}_{\mathfrak{h}_1^k} f, \ldots, \mathscr{R}_{\mathfrak{h}_{d_{n,k}}^k} f\}.$$
   Let $\mathfrak{e}_k := \{\delta_{k,l}\}_{l=1}^{d_{n,k}}$. Then $\{\mathfrak{e}_1^T, \ldots, \mathfrak{e}_{d_{n,k}}^T\}$ is a basis for $\mathbb{R}^{d_{n,k}*}$ and $\mathfrak{e}_\alpha^T(\mathscr{R}f(\,\cdot\,)) = \mathscr{R}_{\mathfrak{h}_\alpha^k} f$ is a partial Higher Riesz transform as is stated in Theorem 8.2.6 iii b).

   Let $\rho \in \mathrm{SO}(n)$, $f \in L^2(\mathbb{R}^n)$. Then
$$\begin{aligned}
\rho\mathscr{R}\rho^{-1} f &= \rho\mathscr{R}f(\rho^{-1}\,\cdot\,) = \rho\big(\mathscr{R}_{\mathfrak{h}_1^k}, \ldots, \mathscr{R}_{\mathfrak{h}_{d_{n,k}}^k}\big) f(\rho^{-1}\,\cdot\,) \\
&= \mathscr{F}^{-1}\Big(\rho\Big(\frac{i\mathfrak{h}_1^k}{\|\cdot\|^n}, \ldots, \frac{i\mathfrak{h}_{d_{n,k}}^k}{\|\cdot\|^n}\Big)\widehat{f}(\rho^{-1})\Big) \\
&= \mathscr{F}^{-1}\Big(\Big(\frac{i\mathfrak{h}_1^k(\rho\,\cdot\,)}{\|\cdot\|^n}, \ldots, \frac{i\mathfrak{h}_{d_{n,k}}^k(\rho\,\cdot\,)}{\|\cdot\|^n}\Big)\widehat{f}\Big) \\
&= \mathscr{F}^{-1}\Big(\Big(\frac{i\mathscr{T}_\rho\mathfrak{h}_1^k}{\|\cdot\|^n}, \ldots, \frac{i\mathscr{T}_\rho\mathfrak{h}_{d_{n,k}}^k}{\|\cdot\|^n}\Big)\widehat{f}\Big) = \big(\mathscr{R}_{\mathscr{T}_\rho\mathfrak{h}_1^k}, \ldots, \mathscr{R}_{T_\rho\mathfrak{h}_{d_{n,k}}^k}\big) f \\
&= S_\rho \mathscr{R} f
\end{aligned}$$
   That is $\big(S_\rho, \mathscr{R}(L^2(\mathbb{R}^n))\big)$ is equivalent to the representation $(\mathscr{T}_\rho, \mathfrak{H}_k(\mathbb{R}^n))$.

   Since S_ρ is a linear combination of the elements of the vector $\mathscr{R}f$, it can be considered as a mapping in $V = \mathbb{R}^{d_{n,k}}$. As such $S_\rho = \mathfrak{D}_\rho^k$.

3. Let $p = \sum_{l=1}^{d_{n,k}} a_l \mathfrak{h}_l^k \in \mathfrak{H}_k$. Note that $\mathscr{R}_p|_{H_0} : H_0(\mathbb{R}^n) \to H_k(\mathbb{R}^n)$.

   Let $f \in H_k(\mathbb{R}^n)$. Then there exist $f_1, \ldots, f_{d_{n,k}} \in H_0(\mathbb{R}^n)$ and $a = (a_1, \ldots, a_{d_{n,k}}) \in \mathbb{R}^{d_{n,k}}$ such that
$$f = \sum_{m=1}^{d_{n,k}} a_m \mathscr{R}_{\mathfrak{h}_m^k} f_m.$$
   Hence a representation of the rotation group is given by
$$\mathscr{S}_\rho : H_k \to H_k,\ f \mapsto S_\rho(f) := \sum_{m=1}^{d_{n,k}} (\mathfrak{D}_\rho^k a)_m \mathscr{R}_{\mathfrak{h}_m^k} f_m.$$
   The representation $\big(S_\rho, H_k(\mathbb{R}^n)\big)$ is equivalent to the representation $(\mathscr{T}_\rho, \mathfrak{H}_k(\mathbb{R}^n))$.

8.2.2 Higher Riesz transforms of frames

Using the proper normalization, the higher Riesz transform of degree k is a unitary operator on the Clifford-Hilbert module $L^2(\mathbb{R}^n)_{d_{n,k}}$. As a consequence the higher Riesz transform of the mother wavelet of a wavelet frame for $L^2(\mathbb{R}^n)$ is the set of mother wavelets for a multiwavelet frame of $L^2(\mathbb{R}^n)$. This multiwavelet frame is steerable with respect to rotation iff the original wavelet frame was derived from a radial, i.e. rotation invariant mother wavelet.

To proof this we will first proof that the hypercomplex higher Riesz transforms are unitary operators.

Theorem 8.2.7 (Unitarity of hypercomplex higher Riesz transforms)
LET $\{\mathfrak{h}_1^k, \dots, \mathfrak{h}_{d_{n,k}}^k\} \subset \mathfrak{H}_k$ *be a basis of* $\mathfrak{H}_k$, *such that* $\{\mathfrak{h}_l^k|_{S^{n-1}}\}_l$ *are orthogonal with respect to* $\langle \cdot , \cdot \rangle_{L^2(S^{n-1})}$ *and normalized such that for all* $\omega \in S^{n-1}$

$$\sum_{l=1}^{d_{n,k}} \left(\mathfrak{h}_l^k(x)\right)^2 = d_{n,k}|x|^{2k},\ \forall x \in \mathbb{R}^n. \tag{8.2}$$

Furthermore, let

$$\mathscr{R} : L^2(\mathbb{R}^n)_{d_{n,k}} \to L^2(\mathbb{R}^n)_{d_{n,k}},\ f \mapsto d_{n,k}^{-1/2} \sum_{\boldsymbol{\beta} \in \mathscr{O}_{d_{n,k}}} \sum_{\alpha=1}^{d_{n,k}} e_{\boldsymbol{\beta}} e_\alpha \mathscr{R}_{\mathfrak{h}_\alpha^k} f_{\boldsymbol{\beta}}.$$

THEN *for* $k \in 2\mathbb{N} - 1$ *the operator* $\mathscr{R}$ *is unitary and self adjoint in the Clifford-Hilbert module* $L^2(\mathbb{R}^n)_{d_{n,k}}$.
For $k \in 2\mathbb{N}$ *the operator* $i\mathscr{R}$ *is self adjoint and* $\mathscr{R}$ *is unitary in the Clifford-Hilbert module* $L^2(\mathbb{R}^n)_{d_{n,k}}$.

Proof. (8.1) in Theorem 8.1.8 on page 140 states that we can indeed normalize any basis of spherical harmonics whose elements are orthogonal to each other such that (8.2) is fulfilled.

Since the polynomials $\mathfrak{h}_m^k$ are k-homogenous, it follows that

$$\mathfrak{h}_m^k(x) = \|x\|^k \mathfrak{h}_m^k(\omega),\ \forall x = \omega\|x\| \in \mathbb{R}^n,\ \text{where } \omega = \frac{x}{\|x\|} \in S^{n-1}.$$

Hence

$$\sum_{m=1}^{d_{n,k}} \left(\mathfrak{h}_m^k(x)\right)^2 = \|x\|^{2k},\ \forall x \in \mathbb{R}^n. \tag{8.3}$$

Recall from Theorem 8.2.2 that

$$\widehat{\mathscr{R}_{\mathfrak{h}_\alpha^k} f}(\xi) = m_p(\xi)\widehat{f}(\xi) := i^k \frac{\mathfrak{h}_\alpha^k(\xi)}{|\xi|^k} \widehat{f}(\xi).$$

Let $k \in 2\mathbb{N}-1$. Then $\mathscr{R}$ is self adjoint, since

$$\begin{aligned}
\langle\!\langle \mathscr{R}f, g\rangle\!\rangle = \langle\!\langle \widehat{\mathscr{R}f}, \widehat{g}\rangle\!\rangle &= \sum_{\alpha=1}^{d_{n,k}} \sum_{\boldsymbol{\beta},\boldsymbol{\gamma}\in\mathscr{O}_{d_{n,k}}} e_{\boldsymbol{\beta}} e_\alpha \overline{e_{\boldsymbol{\gamma}}} \int_{\mathbb{R}^n} \frac{(-1)^{(k-1)/2} i \mathfrak{h}_\alpha^k(x)}{|x|^k} \widehat{f_{\boldsymbol{\beta}}}(x) \overline{\widehat{g_{\boldsymbol{\gamma}}}}(x)\,dx \\
&= \sum_{\alpha=1}^{d_{n,k}} \sum_{\boldsymbol{\beta},\boldsymbol{\gamma}\in\mathscr{O}_{d_{n,k}}} e_{\boldsymbol{\beta}} (-\overline{e_\alpha}) \overline{e_{\boldsymbol{\gamma}}} \int_{\mathbb{R}^n} \widehat{f_{\boldsymbol{\beta}}}(x) \overline{\frac{-(-1)^{(k-1)/2} i \mathfrak{h}_\alpha^k(x)}{|x|^k} \widehat{g_{\boldsymbol{\gamma}}}}(x)\,dx \\
&= \sum_{\alpha=1}^{d_{n,k}} \sum_{\boldsymbol{\beta},\boldsymbol{\gamma}\in\mathscr{O}_{d_{n,k}}} e_{\boldsymbol{\beta}} \overline{e_{\boldsymbol{\gamma}} e_\alpha} \int_{\mathbb{R}^n} \widehat{f_{\boldsymbol{\beta}}}(x) \overline{\frac{(-1)^{(k-1)/2} i \mathfrak{h}_\alpha^k(x)}{|x|^k} \widehat{g_{\boldsymbol{\gamma}}}(x)}\,dx \\
&= \langle\!\langle \widehat{f}, \widehat{\mathscr{R}g}\rangle\!\rangle \\
&= \langle\!\langle f, \mathscr{R}g\rangle\!\rangle,
\end{aligned}$$

where we used Plancherel's equality. (See Theorem 3.2.15 on page 58).
Now we need to show that $\langle\!\langle \mathscr{R}f, \mathscr{R}g\rangle\!\rangle = \langle\!\langle f, g\rangle\!\rangle$. Indeed

$$\begin{aligned}
\langle\!\langle \mathscr{R}f, \mathscr{R}g\rangle\!\rangle = \langle\!\langle \widehat{\mathscr{R}\mathscr{R}f}, \widehat{g}\rangle\!\rangle &= \sum_{\beta,\gamma=1}^{d_{n,k}} \sum_{\boldsymbol{\alpha},\boldsymbol{\delta}\in\mathscr{O}_{d_{n,k}}} e_{\boldsymbol{\alpha}} e_\gamma e_\beta \overline{e_{\boldsymbol{\delta}}} \int_{\mathbb{R}^n} \frac{-\mathfrak{h}_\beta^k(\xi)\mathfrak{h}_\gamma^k(\xi)}{\|\xi\|^{2k}} \widehat{f_{\boldsymbol{\alpha}}}(\xi) \overline{\widehat{g_{\boldsymbol{\delta}}}}(\xi)\,d\xi \\
&= \sum_{\boldsymbol{\alpha},\boldsymbol{\delta}\in\mathscr{O}_{d_{n,k}}} e_{\boldsymbol{\alpha}} \Big(\sum_{1\le\beta<\gamma\le d_{n,k}} \underbrace{(e_\gamma e_\beta - e_\beta e_\gamma)}_{=0} \int_{\mathbb{R}^n} \frac{-\mathfrak{h}_\beta^k(\xi)\mathfrak{h}_\gamma^k(\xi)}{\|\xi\|^{2k}} \widehat{f_{\boldsymbol{\alpha}}}(\xi) \overline{\widehat{g_{\boldsymbol{\delta}}}}(\xi)\,d\xi \\
&\quad + \int_{\mathbb{R}^n} \underbrace{\sum_{\beta=1}^{d_{n,k}} \underbrace{e_\beta^2}_{=-1} \frac{-\big(\mathfrak{h}_\beta^k(\xi)\big)^2}{\|\xi\|^{2k}}}_{=1,\text{ by (8.3)}} \widehat{f_{\boldsymbol{\alpha}}}(\xi) \overline{\widehat{g_{\boldsymbol{\delta}}}}(\xi)\,d\xi \Big) \overline{e_{\boldsymbol{\delta}}} = \langle\!\langle \widehat{f}, \widehat{g}\rangle\!\rangle \\
&= \langle\!\langle f, g\rangle\!\rangle.
\end{aligned}$$

The proof for $k \in 2\mathbb{N}$ is quite similar. Since in this case $i^k \in \mathbb{R}$ we have to multiply by i to use the argumentation above. This proves that $i\mathscr{R}$ is self adjoint. It follows that $\mathscr{R}^* = -\mathscr{R}$. This yields the negative sign needed in the proof of unitarity. □

Using Theorem 8.2.7, we can show that the higher Riesz transforms map Clifford frames onto Clifford frames.

Theorem 8.2.8 (Frame property of higher Riesz transforms)
LET $\{f_l\}_{l\in\mathbb{N}} \subset L^2(\mathbb{R}^n,\mathbb{R})$ *be a frame for* $L^2(\mathbb{R}^n,\mathbb{R})$ *with frame bounds* A *and* B. *Furthermore let* $\{\mathfrak{h}_1^k,\ldots,\mathfrak{h}_{d_{n,k}}^k\} \subset \mathfrak{H}_k$ *be a basis of* $\mathfrak{H}_k$ *as in Theorem 8.2.7 and let* $\mathscr{R}$ *be the corresponding hypercomplex higher Riesz transform.*

(i) Then $\{f_l\}_{l\in\mathbb{N}}$ *is a Clifford frame for* $L^2(\mathbb{R}^n,\mathbb{R})_{d_{n,k}} \cong L^2(\mathbb{R}^n,\mathbb{R}_{d_{n,k}})$ *with the same frame bounds* A *and* B.

(ii) $\{\mathscr{R}f_l\}_{l\in\mathbb{N}}$ *is a Clifford frame for* $L^2(\mathbb{R}^n,\mathbb{R})_{d_{n,k}}$ *with the same frame bounds* A *and* B.

Proof. The proof follows directly from Theorem 5.2.13 on page 86 just as the proof of Theorem 5.3.3. □

8.2.3 A phase concept for higher Riesz transforms

In Definition 8.2.4 we have established hypercomplex higher Riesz transforms. In this section we define a higher monogenic signal to obtain a phase amplitude decomposition of a signal. Furthermore we present a kind of Cauchy-Riemann equation which is satisfied by the Poisson transform of the higher monogenic signal just as the Poisson transform of the monogenic signal satisfies the Cauchy Riemann equations (3.4). (See Theorem 4.5.2 on page 70.)

The first step is to define a phase direction for the higher Riesz transform of a function.

Definition 8.2.9 (Phase direction for higher Riesz transforms)
LET $f \in L^2(\mathbb{R}^n)$ *and let* $\left(\mathscr{R}_{\mathfrak{h}_l^k}\right)_{k=1}^{d_{n,k}} \in L^2(\mathbb{R}^n)_{d_{n,k}}$ *be its higher Riesz transform.*

THEN *the* **phase direction** d *is defined as*

$$d : \mathbb{R}^n \to \mathbb{R}^{d_{n,k}},\ d = \frac{\mathscr{R}_k f}{|\mathscr{R}_k f|} = \frac{\left(\mathscr{R}_{\mathfrak{h}_l^k} f\right)_{k=1}^{d_{n,k}}}{|\mathscr{R}_k f|}.$$

Remark 8.2.10 (Interpretation of the phase direction)
The phase direction of a higher Riesz transform is at first view hard to interpret, since the domain of the signal is n-dimensional while the phase direction is $d_{n,k}$ dimensional. It seems reasonable to interpret the phase by the spherical harmonic $d = \sum_{l=1}^{d_{n,k}} d_l \mathfrak{h}_l^k$. In the case of the Riesz transforms, that is in the case that $k = 1$, this polynomial has exactly one preferred direction - the phase direction of the Riesz transform.

For $n = 2$ the higher Riesz transform has exactly k preferred directions which are equally spaced. It can easily be seen in the explicit formulas in Theorem 8.3.1 and the corresponding Figure 8.1 that directions which defer by a multiple of $\frac{\pi}{k}$ can not be distinguished. This results in a directional ambiguity.

In the case $n = 3$, $k > 1$ a finer distinction is possible - there exist spherical harmonics with exactly one preferred direction and spherical harmonics with more then one preferred direction. (See Figure 8.2 and Figure 8.3.)

Following the definition of monogenic phase in section 4.4 we can establish phase and amplitude for higher Riesz transforms:

Definition 8.2.11 (Amplitude - phase decomposition of higher Riesz transforms)
LET $f \in L^2(\mathbb{R}^n)$.

THEN *the* **higher monogenic signal** *is defined as*

$$f_{m,k} := f + \mathscr{R}_k f \in L^2(\mathbb{R}^n, \mathbb{R}^{d_{n,k}+1}).$$

We can decompose $f_{m,k}$ in the same way as the monogenic signal:

$$\begin{aligned} f_{m,k} &= |f_{m,k}| \left(\frac{f}{|f_{m,k}|} + \frac{\mathscr{R}_k f}{|\mathscr{R}_k f|} \frac{|\mathscr{R}_k f|}{|f_{m,k}|} \right) \\ &= a\Big(\cos(\phi) + d \sin(\phi)\Big) = a \exp(d\phi), \end{aligned} \tag{8.4}$$

where $a = |f_{m,k}|$ is called the **amplitude**, $\phi = \arg\left(\frac{f}{|f_{m,k}|} + i\frac{|\mathscr{R}_k f|}{|f_{m,k}|}\right)$ *is called the* **phase** *and $d = \frac{\mathscr{R}_k f}{|\mathscr{R}_k f|}$ is called the* **phase direction** *of the signal f.*

Remark 8.2.12 (Higher Riesz transforms of wavelets)
Let $\psi \in L^2(\mathbb{R}^n)$ such that ψ is a mother wavelet for a wavelet frame $\{D^j T_t \psi\}_{j\in\mathbb{Z}, t\in\mathbb{Z}^n}$ of $L^2(\mathbb{R}^n)$. Let $f \in L^2(\mathbb{R}^n)$.

- *Then by Theorem 8.2.8 it follows that $\{D^j T_t \mathscr{R}\psi\}_{j\in\mathbb{Z}, t\in\mathbb{Z}^n}$ is a multi wavelet frame of $L^2(\mathbb{R}^n)$. By Example 8.2.1 3. this frame is steerable.*

- *The phase direction defined in Definition 8.2.9, amplitude and phase defined in Definition 8.2.11 are defined in the same way for the coefficients of this wavelet decomposition.*

$$W_{\psi_{m,k}}(f) = \{\langle\langle f, D^j T_t \psi_{m,k}\rangle\rangle\}_{j\in\mathbb{Z}, t\in\mathbb{Z}^n} = \{a_{j,t} \exp(d_{j,t}\phi_{j,t})\}_{j\in\mathbb{Z}, t\in\mathbb{Z}^n}.$$

- *Let $\psi \in \mathfrak{H}_0$. Then $\{\mathscr{R}_{\mathfrak{h}_l^k}\psi\}_{l=1}^{d_{n,k}}$ is a basis for a linear subspace of H_k. The phase direction of the wavelet transform corresponds to the element*

$$\mathscr{R}_{d_{j,t}}\psi \subset \operatorname{span}\{\mathscr{R}_{\mathfrak{h}_l^k}\psi,\ l = 1,\dots,d_{n,k}\} \subset \mathfrak{H}_k$$

  of norm $\|\mathscr{R}_{d_{j,t}}\psi\| = \|\psi\|$ that has the highest scalar product with $T_{-t}D^{-j}f$.

  That is we approximate the signal f with the optimal element $R_{d_{j,t}}\psi$ in the subspace $\operatorname{span}\{\mathscr{R}_{\mathfrak{h}_l^k}\psi\}_{l=1,\dots,d_{n,k}}$ of $\mathfrak{H}_k(\mathbb{R}^n)$. To find this element we need the higher Riesz transform which consists of a basis of this subspace.

To state the Cauchy Riemann equation that the Poisson transform of the higher Riesz transforms satisfy, we first define the Cauchy-Riemann operators whose kernel satisfies the generalized Cauchy Riemann equations.

Definition 8.2.13 (Higher Dirac operator)
LET *$f \in W^{k,2}(\mathbb{R}^n)$. Furthermore, let $\{\mathfrak{h}_l^k\}_{l=1}^{d_{n,k}}$ be a ONB of $\mathfrak{H}_k$. Then*

$$\mathfrak{h}_l^k = \sum_{\substack{\boldsymbol{\beta}\in\mathbb{N}_0^n \\ |\boldsymbol{\beta}|\le k}} r_{l,\boldsymbol{\beta}}^k x^{\boldsymbol{\beta}},$$

for some $r_{l,\boldsymbol{\beta}}^k \in \mathbb{C}$.

THEN *the* **higher Dirac operator** *of order k is defined as*

$$D_k : W_2^k(\mathbb{R}^n, \mathbb{R}_{d_{n,k}}) \to L^2(\mathbb{R}^n, \mathbb{R}_{d_{n,k}}),\ f \mapsto \sum_{\alpha=1}^{d_{n,k}} e_\alpha d_{n,k}^{-1/2} \mathfrak{h}_\alpha^k(\partial) f,$$

where $\mathfrak{h}_\alpha^k(\partial) := \sum_{|\boldsymbol{\beta}|\le k} r_{\boldsymbol{\beta},\alpha}^k \partial^{\boldsymbol{\beta}} = \sum_{|\boldsymbol{\beta}|\le k} r_{\boldsymbol{\beta},\alpha}^k \frac{\partial^{\beta_1}}{\partial x_1^{\beta_1}} \cdots \frac{\partial^{\beta_n}}{\partial x_n^{\beta_n}}$.

The corresponding **higher Cauchy-Riemann operator** *is*

$$\partial_k := \frac{\partial^k}{\partial x_0^k} + D_k : W^{k,2}(\mathbb{R}^{n+1}, \mathbb{R}_{d_{n,k}}) \to L^2(\mathbb{R}^{n+1}, \mathbb{R}_{d_{n,k}}).$$

The **conjugate Cauchy-Riemann operator** *is $\underline{\partial_k} := \frac{\partial^k}{\partial x_0^k} - D_k$*

Corollary 8.2.14 (Factorization of powers of the Laplacian)
LET $\triangle = \sum_{\alpha=1}^{d_{n,k}} \partial_\alpha^2 : W^{2,2}(\mathbb{R}^n) \to L^2(\mathbb{R}^n)$ *be the Laplacian and let*

$$\triangle^k : W^{2k,2}(\mathbb{R}^n) \to L^2(\mathbb{R}^n),\ f \mapsto (\prod_{l=1}^{k} \triangle) f$$

denote its k-th power.
THEN

$$D_k^2 = -\triangle^k.$$

Furthermore

$$\partial_k \underline{\partial_k} = \frac{\partial^{2k}}{\partial x_0^{2k}} + \triangle^k.$$

Proof. Let $f \in W^{2k,2}(\mathbb{R}^n)$.

$$\begin{aligned}
\mathscr{F}(D_k^2 f)(\xi) &= \mathscr{F}\left(\sum_{\alpha=1}^{d_{n,k}} \sum_{\beta=1}^{d_{n,k}} e_\alpha e_\beta d_{n,k}^{-1} \mathfrak{h}_\alpha^k(\partial) \mathfrak{h}_\beta^k(\partial) f\right)(\xi) = d_{n,k}^{-1} \sum_{\alpha=1}^{d_{n,k}} \sum_{\beta=1}^{d_{n,k}} e_\alpha e_\beta \mathfrak{h}_\alpha^k(-2\pi i \xi) \mathfrak{h}_\beta^k(-2\pi i \xi) \widehat{f}(\xi) \\
&= (-2\pi i)^{2k} d_{n,k}^{-1} \Big(-\sum_{\alpha=1}^{d_{n,k}} \big(\mathfrak{h}_\alpha^k(\xi)\big)^2 + \sum_{\substack{\alpha,\beta=1,\dots,d_{n,k} \\ \alpha<\beta}} e_\alpha e_\beta \mathfrak{h}_\alpha^k(\xi)\mathfrak{h}_\beta^k(\xi) + e_\beta e_\alpha \mathfrak{h}_\beta^k(\xi) \mathfrak{h}_\alpha^k(\xi)\Big) \widehat{f}(\xi) \\
&= (-2\pi i)^{2k} d_{n,k}^{-1} \Big(-d_{n,k} |\xi|^{2k} + \sum_{\substack{\alpha,\beta=1,\dots,d_{n,k} \\ \alpha<\beta}} e_\alpha e_\beta \underbrace{\big(\mathfrak{h}_\alpha^k(\xi)\mathfrak{h}_\beta^k(\xi) - (\xi)\mathfrak{h}_\alpha^k(\xi)\mathfrak{h}_\beta^k\big)}_{=0}\Big) \widehat{f}(\xi) \\
&= -\mathscr{F}\big(\triangle^k f\big)(\xi), \text{ f.a.a. } \xi \in \mathbb{R}^n.
\end{aligned}$$

$$\partial_k \underline{\partial_k} f(x) = \frac{\partial^2}{\partial x_0^2} f(x) + \big(\frac{\partial}{\partial x_0} D_k - D_k \frac{\partial}{\partial x_0}\big) f(x) - D_k^2 f(x),\ \forall f \in W^{2k,2}(\mathbb{R}^{n+1}), \text{ f.a.a. } x \in \mathbb{R}.$$

□

Theorem 8.2.15
LET

$$f = \sum_{\alpha=0}^{d_{n,k}} e_\alpha f_\alpha \in L^p(\mathbb{R}^n, \mathbb{R}_+^{d_{n,k}+1}),\ 1 < p < \infty, \qquad \text{if } k \text{ is odd,}$$

and

$$f = e_0 f_0 - i \sum_{\alpha=1}^{d_{n,k}} e_\alpha f_\alpha \in L^p(\mathbb{R}^n, \mathbb{R}_+^{d_{n,k}+1}),\ 1 < p < \infty, \qquad \text{if } k \text{ is even.}$$

Furthermore, let

$$u_\alpha(x, x_0) := P_{x_0} * f_\alpha(x),\ \forall \alpha = 0, \dots, d_{n,k}, x_0 \in \mathbb{R}^+,\ x \in \mathbb{R}^n.$$

THEN $f_\alpha = \mathscr{R}_{\mathfrak{h}_\alpha^k}(f_0)$, $\alpha = 1, \dots, d_{n,k}$, *iff u satisfies*

$$\partial_k u = 0, \text{ if } k \text{ is odd, and}$$

$$\big(\frac{\partial^k}{\partial x_0^k} + iD\big) u = 0, \text{ if } k \text{ is even.}$$

This is equivalent to the following set of **generalized Cauchy Riemann equations***:*

$$\mathfrak{h}^k_\beta(D)u_{\boldsymbol{\alpha}}(x) - \mathfrak{h}^k_\alpha(D)u_\beta(x) = 0, \qquad \forall \alpha,\beta = 1,\dots,d_{n,k}; \tag{8.5}$$

$$\frac{\partial^k u_\alpha}{\partial x_0^k}(x) + (-1)^{k+1}\mathfrak{h}^k_\alpha(D)u_0(x) = 0, \qquad \forall \alpha = 1,\dots,d_{n,k}; \tag{8.6}$$

$$\sum_{\alpha=1}^{n}\mathfrak{h}^k_\alpha(D)u_\alpha(x) - \frac{\partial^k u_0}{\partial x_0^k}(x) = 0, \qquad \text{f.a.a. } x_0 \in \mathbb{R}^+,\ x \in \mathbb{R}^n. \tag{8.7}$$

Proof. The equalities in this proof hold for almost all $x \in \mathbb{R}^n$, $x_0 \in \mathbb{R}^+$.
If k is odd, then

$$\begin{aligned}\partial_k u(x,x_0) = &\frac{\partial^k}{\partial x_0^k}u_0(x,x_0) - \sum_{\alpha=1}^{d_{n,k}}\mathfrak{h}^k_\alpha(D)u_\alpha(x,x_0)\\ &+\sum_{\alpha=1}^{d_{n,k}}e_\alpha\big(\mathfrak{h}^k_\alpha(D)u_0(x,x_0) + \frac{\partial^k}{\partial x_0^k}u_\alpha(x,x_0)\big)\\ &+\sum_{\alpha\neq\beta=1}^{d_{n,k}}e_\alpha e_\beta\big(\mathfrak{h}^k_\alpha(D)u_\beta(x,x_0) - \mathfrak{h}^k_\beta(D)u_\alpha(x,x_0))\big).\end{aligned}$$

This equals 0 iff the generalized Cauchy Riemann equations (8.5), (8.6) and (8.7) hold.
If k is even, then

$$\begin{aligned}\big(\frac{\partial^k}{\partial x_0^k} + iD\big)u(x,x_0) = &\frac{\partial^k}{\partial x_0^k}u_0(x,x_0) - i(-i)\sum_{\alpha=1}^{d_{n,k}}\mathfrak{h}^k_\alpha(D)u_\alpha(x,x_0)\\ &+\sum_{\alpha=1}^{d_{n,k}}e_\alpha\big(i\mathfrak{h}^k_\alpha(D)u_0(x,x_0) - i\frac{\partial^k}{\partial x_0^k}u_\alpha(x,x_0)\big)\\ &+\sum_{\alpha\neq\beta=1}^{d_{n,k}}i(-i)e_\alpha e_\beta\big(\mathfrak{h}^k_\alpha(D)u_\beta(x,x_0) - \mathfrak{h}^k_\beta(D)u_\alpha(x,x_0))\big).\end{aligned}$$

Which is equal to 0 iff the generalized Cauchy Riemann (8.5), (8.6) and (8.7) equations hold.
Suppose $f_\alpha = R_{\mathfrak{h}^k_\alpha} f_0$, then $\widehat{f_\alpha}(t) = \frac{i^k\mathfrak{h}^k_\alpha(t)}{|t|^k}\widehat{f_0}(t)$ and, hence,

$$u_\alpha(x,x_0) = \int_{\mathbb{R}^n}\widehat{f_0}(t)\frac{\mathfrak{h}^k_\alpha(it)}{|t|^k}e^{-2\pi|t|x_0}e^{2\pi itx}dt,\ \forall\alpha = 1,\dots,d_{n,k},$$

and

$$u_0(x,x_0) = \int_{\mathbb{R}^n}\widehat{f_0}(t)e^{-2\pi|t|x_0}e^{2\pi itx}dt.$$

By the dominated convergence theorem [41], we may differentiate under the integral sign, and hence obtain

$$\frac{\partial^k u_0}{\partial x_0^k}(x,x_0) = (-2\pi)^k \int_{\mathbb{R}^n} \widehat{f_0}(t)|t|^k e^{-2\pi|t|x_0} e^{2\pi i t x} dt$$
$$\mathfrak{h}_\beta^k(\partial) u_0(x,x_0) = \int_{\mathbb{R}^n} \widehat{f_0}(t)(2\pi i)^k \mathfrak{h}_\beta^k(t) e^{-2\pi|t|x_0} e^{2\pi i t x} dt$$
$$\frac{\partial^k u_\beta}{\partial^k x_0}(x,x_0) = \int_{\mathbb{R}^n} \widehat{f_0}(t)(-2\pi)^k i^k \mathfrak{h}_\beta^k(t) e^{-2\pi|t|x_0} e^{2\pi i t x} dt$$
$$\mathfrak{h}_\beta^k(\partial)\partial u_\alpha(x,x_0) = \int_{\mathbb{R}^n} \widehat{f_0}(t)(2\pi i)^k i^k \frac{\mathfrak{h}_\alpha^k(t)\mathfrak{h}_\beta^k(t)}{|t|^k} e^{-2\pi|t|x_0} e^{2\pi i t x} dt$$

Now (3.4) is easy to check. The first equation (8.5) follows from

$$\begin{aligned}\mathfrak{h}_\beta^k(\partial)\partial u_\alpha(x,x_0) &= (-2\pi)^k \int_{\mathbb{R}^n} \widehat{f}(t) \frac{\mathfrak{h}_\alpha^k(t)\mathfrak{h}_\beta^k(t)}{|t|^k} e^{-2\pi|t|x_0} e^{2\pi i t x} dt \\ &= \mathfrak{h}_\alpha^k(\partial) u_\beta(x,x_0).\end{aligned}$$

The second identity (8.6) follows from

$$\begin{aligned}\mathfrak{h}_\beta^k(\partial)\partial u_0(x,x_0) &= (2\pi i)^k \int_{\mathbb{R}^n} \widehat{f}(t)\mathfrak{h}_\beta^k(t) e^{-2\pi|t|x_0} e^{2\pi i t x} dt \\ &= (-1)^k \frac{\partial^k u_\beta}{\partial^k x_0}(x,x_0).\end{aligned}$$

Finally the last identity (8.7) follows from

$$\begin{aligned}\sum_{\alpha=1}^{d_{n,k}} \mathfrak{h}_\alpha^k(\partial) u_\alpha(x,x_0) &= (-1)^k \sum_{\alpha=1}^{d_{n,k}} (2\pi)^k \int_{\mathbb{R}^n} \widehat{f}(t) \frac{\left(\mathfrak{h}_\alpha^k(t)\right)^2}{|t|^k} e^{-2\pi|t|x_0} e^{2\pi i t x} dt \\ &= (-2\pi)^k \int_{\mathbb{R}^n} \widehat{f}(t) \frac{\sum_{\alpha=1}^{d_{n,k}} \left(\mathfrak{h}_\alpha^k(t)\right)^2}{|t|^k} e^{-2\pi|t|x_0} e^{2\pi i t x} dt \\ &= (-2\pi)^k \int_{\mathbb{R}^n} \widehat{f}(t)|t|^k e^{-2\pi|t|x_0} e^{2\pi i t x} dt \\ &= \frac{\partial u}{\partial x_0}(x,x_0).\end{aligned}$$

Conversely, let $\beta \in \{1,\dots,d_{n,k}\}$ and $u_\beta(x,x_0) = \int_{\mathbb{R}^n} \widehat{f_\beta}(t) e^{-2\pi|t|x_0} e^{2\pi i t x} dt$. The fact that $\mathfrak{h}_\beta^k(\partial) u_0 = (-1)^k \frac{\partial^k u_\beta}{\partial^k x_0}$, shows that

$$(2\pi i)^k \int_{\mathbb{R}^n} \widehat{f}(t)\mathfrak{h}_\beta^k(t) e^{-2\pi|t|x_0} e^{2\pi i t x} dt = 2\pi \int_{\mathbb{R}^n} \widehat{f_\beta}(t)|t|^k e^{-2\pi|t|x_0} e^{2\pi i t x} dt.$$

Therefore $\widehat{f_\beta}(t) = \frac{i^k \mathfrak{h}_\beta^k(t)}{|t|^k} \widehat{f}(t)$, and so

$$f_\beta = \mathscr{R}_{\mathfrak{h}_\beta^k}(f),\ \beta = 1,\dots,d_{n,k}.$$

□

8.2.4 Composition of higher Riesz transforms

As is easily seen by Figure 8.1, Figure 8.2 and Figure 8.3 the higher Riesz transforms based on spherical harmonics are restricted with respect to their geometry. It might for example be interesting to look at the higher Riesz transforms given by the Fourier multiplier $\left(\frac{ix_l}{|x|}\right)^k$. However if $k \geq 2$ the polynomial $x_l^k|_{S^{n-1}}$ lies in $\mathfrak{P}_k|_{S^{n-1}} = \sum_{l=0}^{\lfloor k/2 \rfloor} \mathscr{H}_{k-2l}$ but not in $\mathscr{H}_k$. In this section we derive higher Riesz transforms based on polynomial spaces composed of spherical harmonics.

The higher Riesz transforms are a vector consisting of the basis elements of a rotation invariant vector space

$$\mathscr{R}_{\mathfrak{H}_k} := \operatorname{span}\{\mathscr{R}_{\mathfrak{h}_l^k},\ l = 1, \ldots, d_{n,k}\}$$

of operators on $L^p(\mathbb{R}^n)$, $1 < p < \infty$. This vector space has an irreducible representation of the rotation group equivalent to the Wigner-$\mathfrak{D}$ matrices. We can compose these spaces to gain new spaces of operators that feature a reducible representation of the rotation group. In fact all such operator spaces that are compatible with wavelets can be composed of such spaces of higher Riesz transforms.

The proof that an operator $\mathscr{T}$ on a certain Clifford-Hilbert module maps frames onto frames is based on Theorem 5.2.13 and depends only on the fact that the operator $\mathscr{T}$ is invertible. If furthermore $\|\mathscr{T}^\dagger\| = \|\mathscr{T}\| = 1$, where $T^\dagger$ is the pseudo-inverse of T, then $\mathscr{T}$ maps tight frames onto tight frames. In Theorem 8.2.8 we chose the higher Riesz transforms $\mathscr{R}_k$ to be unitary. The proof of the unitarity Theorem 8.2.7 depends only on (8.2), namely that

$$\sum_{l=1}^{d_{n,k}} \left(d_{n,k}^{-1/2} \mathfrak{h}_l^k(\omega)\right)^2 = 1,\ \forall \omega \in S^{n-1}.$$

Furthermore, the higher Riesz transforms satisfy $\mathscr{R}_k^{-1} = \mathscr{R}_k^* = (-1)^k \mathscr{R}_k$.

Definition 8.2.16
Let $K = \{k_1, \ldots, k_{\mathfrak{k}}\} \subset \mathbb{N}_0$, where $\mathfrak{k} \in \mathbb{N}$ and $k_l \neq k_m, \forall l \neq m \in K$. We define a higher Riesz transform that consists of basis elements of the spaces $\mathscr{R}_{\mathfrak{H}_{k_r}}$, $r \in K$. Let

$$\begin{aligned}\mathscr{R}_K &: L^2(\mathbb{R}^n)_{\sum_{k\in K} d_{n,k}} \to L^2(\mathbb{R}^n)_{\sum_{k\in K} d_{n,k}},\\ \mathscr{R}_K &:= \sum_{k=k_1,\ldots,k_{\mathfrak{k}}} \sum_{l=1}^{d_{n,k}} e_{\alpha_{k,l}} d_K \mathscr{R}_{\mathfrak{h}_l^k},\end{aligned} \tag{8.8}$$

where $\alpha_{k,l} = \sum_{m=1}^{k-1} d_{n,m} + l$, and $d_K \in \mathbb{C}$ is chosen such that

$$\sum_{k\in K} \sum_{l=1}^{d_{n,k}} \left(d_K \mathfrak{h}_l^k(\omega)\right)^2 = 1,\ \forall \omega \in S^{n-1}. \tag{8.9}$$

Corollary 8.2.17
(8.9) *is easily satisfied by setting*

$$d_K = \Big(\sum_{k\in K} d_{n,k}\Big)^{-1/2}.$$

Proof. Indeed is true that $\sum_{k\in K} \sum_{l=1}^{d_{n,k}} \left(d_K \mathfrak{h}_l^k\right)^2 = d_K^2 \sum_{k\in K} d_{n,k} = 1$. □

Corollary 8.2.18 (Composition of higher Riesz transforms)
Let K as above. Then $\mathscr{R}_{\bigoplus_{k\in K}\mathfrak{H}_k} = \bigoplus_{k\in K}\mathscr{R}_{\mathfrak{H}_k}$ is a rotation invariant vector space of linear bounded operators in $L^p(\mathbb{R}^n)$ invariant under translation and dilation. Its elements are uniquely defined by a vector $a = (a_l^m)_{\substack{l=1,\dots,d_{n,k}\\ k\in K}}$, $a_l^k \in \mathbb{C}$, via

$$\mathscr{R}_a = \sum_{k\in K}\sum_{l=1}^{d_{n,k}} a_l^k \mathscr{R}_{h_l^k}.$$

A reducible representation $\mathscr{S}$ of the rotation group on $\mathscr{R}_{\bigoplus_{k\in K}\mathfrak{H}_k}$ is given by

$$\rho^{-1}\circ\mathscr{R}_a\circ\rho = \mathscr{S}_\rho\mathscr{R}_a := \sum_{k\in K}\sum_{l=1}^{d_{n,k}} (\mathfrak{D}_\rho^k a^k)_l \mathscr{R}_{h_l^k},$$

where $\mathfrak{D}_\rho^l$ is the Wigner-$\mathfrak{D}$ matrix and $a^k \in \mathbb{C}^{d_{n,k}} = (a_1^k,\dots,a_{d_{n,k}}^k)$.
A higher Riesz transform can be defined as a vector of basis elements of $\mathscr{R}_{\bigoplus_{k\in K}\mathfrak{H}_k}$ the canonical example is

$$\mathscr{R}_K := \left(d_K\mathscr{R}_{\mathfrak{h}_l^k}\right)_{\substack{l=1,\dots,d_{n,k}\\ k\in K}}.$$

This higher Riesz transform is steerable because of the representation $\mathscr{S}$ of the rotation group on $\mathscr{R}_{\bigoplus_{k\in K}\mathfrak{H}_k}$.
The adjoint of the hypercomplex higher Riesz transform (8.8) *is*

$$\mathscr{R}_K^* = \sum_{k=k_1,\dots,k_{\mathfrak{k}}}\sum_{l=1}^{d_{n,k}} (-1)^k e_{\alpha_{k,l}} d_K \mathscr{R}_{\mathfrak{h}_l^k}.$$

The same argumentation as in Theorem 8.2.7 yields that $\mathscr{R}_K$ is a unitary operator, whence it is bijective with inverse $\mathscr{R}^*$ and hence using Theorem 5.2.13 on page 86 it maps frames of $L^2(\mathbb{R}^n)$ to multiframes of $L^2(\mathbb{R}^n)$.

EXAMPLE 8.2.2 (Higher Riesz transforms for the spaces $\mathfrak{P}_k$):
Let $K = \{k-2m\}_{m=0}^{\lfloor k/2\rfloor}$. Then $\mathscr{R}_{\bigoplus_{k\in K}\mathfrak{H}_k} = \mathscr{R}_{\mathfrak{P}_k}$. Indeed, let $p\in\mathfrak{P}_k$. Theorem 8.1.4 implies that $\mathfrak{P}_k = \bigoplus_{m=0}^{\lfloor \frac{k}{2}\rfloor} |\cdot|^{2m}\mathfrak{H}_{k-2m}$. As a consequence there exist a_l^{k-2m}, $l=1,\dots,d_{n,k-2m}$, $m=0,\dots,\lfloor\frac{k}{2}\rfloor$ such that

$$p(x) = \sum_{m=0}^{\lfloor\frac{k}{2}\rfloor}\sum_{l=1}^{d_{n,k-2m}} a_l^{k-2m} h_l^{k-2m}(x)|x|^{2m},\ \forall x\in\mathbb{R}^n.$$

Let $f\in L^2(\mathbb{R}^n)$. Then

$$\begin{aligned}
\widehat{\mathscr{R}_p f}(\xi) &= \frac{i^k p(\xi)}{|\xi|^k}\widehat{f}(\xi) = \frac{i^k\sum_{m=0}^{\lfloor\frac{k}{2}\rfloor}\sum_{l=1}^{d_{n,k-2m}} a_l^{k-2m}\mathfrak{h}_l^{k-2m}(x)|\xi|^{2m}}{|\xi|^k}\widehat{f}(\xi)\\
&\sum_{m=0}^{\lfloor\frac{k}{2}\rfloor}\sum_{l=1}^{d_{n,k-2m}} i^{2m}\frac{i^{k-2m}a_l^{k-2m}\mathfrak{h}_l^{k-2m}(\xi)|\xi|^{2m}}{|\xi|^k}\widehat{f}(\xi)\\
&= \sum_{m=0}^{\lfloor\frac{k}{2}\rfloor}\sum_{l=1}^{d_{n,k-2m}} i^{2m}\frac{i^{k-2m}a_l^{k-2m}\mathfrak{h}_l^{k-2m}(\xi)}{|\xi|^{k-2m}}\widehat{f}(\xi)\\
&= \sum_{m=0}^{\lfloor\frac{k}{2}\rfloor}\sum_{l=1}^{d_{n,k-2m}} (-1)^m a_l^{k-2m}\widehat{\mathscr{R}_{\mathfrak{h}_l^{k-2m}} f}(\xi).
\end{aligned}$$

That is we can write any partial higher Riesz transform with respect to a homogenous polynomial of order k as a linear combination of the the partial higher Riesz transforms of order $k-2m$, $m = 0,\dots,\lfloor\frac{k}{2}\rfloor$.

Let $\mathscr{R}_{\mathfrak{P}_k} := \operatorname{span}\{\mathscr{R}_p : p \in \mathfrak{P}_k\}$. Then $\mathscr{R}_{\mathfrak{P}_{k-2}} \subset \mathscr{R}_{\mathfrak{P}_{k-2}} \oplus \mathscr{R}_{\mathfrak{H}^k} = \mathscr{R}_{\mathfrak{P}_k}$.

In view of Remark 8.2.3 this inclusion is obvious, since on the unit circle the polynomial $p \in \mathfrak{P}_k(\mathbb{R}^n)$ and the polynomial $|\cdot|^{2l} p \in \mathfrak{P}_{k+2l}(\mathbb{R}^n)$ agree. As a consequence the space $\mathscr{R}_{\mathfrak{P}_k}$ is the space of higher Riesz transforms with respect to polynomials of degrees $k-2m$, where $m = 0,\dots,\lfloor\frac{k}{2}\rfloor$.

Let d^k be the phase direction corresponding to the higher Riesz transform of order k.

Then $d(x) := \sum_{l=1}^{d_{n,k}} d_l^k \mathfrak{h}_l^k(x) - \sum_{l=1}^{d_{n,k-2}} d_l^{k-2} \|x\|^2 h_l^{k-2}(x) \in \mathfrak{P}_k$

$$d(x) := \sum_{m=0}^{\lfloor\frac{k}{2}\rfloor} \sum_{l=1}^{d_{n,k-2m}} (-1)^m d_l^{k-2m} h_l^{k-2m}(x) \|x\|^{2m}$$

defines a phase direction in the homogenous polynomials of degree k.

Remark 8.2.19 (Alternative bases for $\mathscr{R}_{\bigoplus_{l=1}^m \mathfrak{H}_{k_l}}$)
Let us state a method to generate alternative bases of $\mathscr{R}_{\bigoplus_{l=1}^m \mathfrak{H}_{k_l}}$ from the standard basis which is given by $\{\mathscr{R}_{h_l^k}\}_{l=1,\dots,d_{n,k},k\in K}$. For some $a = (a_l^k)_{l=1,\dots,d_{n,k};k\in K}$, $a_l^k \in \mathbb{C}$, let

$$a \cdot \mathscr{R}_K := \sum_{k\in K} \sum_{l=1}^{d_{n,k}} a_l^k d_K \mathscr{R}_{\mathfrak{h}_l^k}.$$

The basis vectors of the standard basis can be written as $\mathscr{R}_{h_l^k} = a(l,k) \cdot \mathscr{R}_K$, where

$$a(l,k) = \left(a_m^r(l,k)\right)_{m=1,\dots,d_{n,k},r\in K} = \left(\delta_{l,m}\delta_{k,r}\right)_{m=1,\dots,d_{n,k},r\in K} \in \mathbb{R}^{\sum_{k\in K} d_{n,k}}.$$

Let $A \in U(\sum_{k\in K} d_{n,k})$ be a orthogonal matrix on $\mathbb{R}^{\sum_{k\in K} d_{n,k}}$. It is a simple fact from linear algebra that A maps the orthonormal basis $\left\{d_K \mathscr{R}_{\mathfrak{h}(l,k)}\right\}_{l,k}$ to an orthonormal basis $\left\{(Aa)_l^k \cdot \mathscr{R}_K\right\}_{l,k}$.

Let $B \in Gl(\sum_{k\in K} d_{n,k})$ be a invertible matrix with inverse B^{-1}. Then $\left\{(Ba)_l^k \cdot \mathscr{R}_K\right\}_{l,k}$ is a basis for the space $\mathfrak{H}_K$. Let $\{\psi_r\}_r \subset L^2(\mathbb{R}^n)$ be a frame and let $\{\phi_r\}_r \subset L^2(\mathbb{R}^n)$ be the dual frame. Then the dual frame of the frame

$$\left\{(Ba)_l^k \cdot \mathscr{R}_K \psi_r\right\}_{r;l=1,\dots,d_{n,k},k\in K}$$

is the frame

$$\left\{\left(B^{-1}\left((-1)^j a(m,j)\right)_{m=1,\dots,d_{n,j};j\in K}\right)_l^k \cdot \mathscr{R}_K \psi_r\right\}_{r;l=1,\dots,d_{n,k},k\in K}.$$

Up to now we have shown how to combine a set of higher Riesz transforms based on spherical harmonics into a new higher Riesz transform and how to change the basis elements of the higher Riesz transform. Next we show how to do it the other way round: We chose a polynomial and build a basis for the minimal space of higher Riesz transforms such that the partial higher Riesz transform with respect to the chosen polynomial is an element of the higher Riesz transform.

Remark 8.2.20 (Higher Riesz transforms based on polynomials)
Inspired by the steerable filters of Freeman and Adelson in [18] we want to state another, more

constructive method to construct a steerable set of operators on $L^2(\mathbb{R}^n)$ based on a single operator. For this purpose let p be a polynomial of degree $k \in \mathbb{N}$. We assume that p is normalized such that

$$\int_{S^{n-1}} |p(x)|^2 d\sigma(x) = 1.$$

Then there exist a set $K \subset \mathbb{N}$ and $(p_l^k)_{l,k} \in \mathbb{R}^{\sum_{k\in K} d_{n,k}}$ such that

$$p = \sum_{k\in K} p_l^k \mathfrak{h}_l^k \in \bigoplus_{k\in K} \mathfrak{H}_k.$$

By our assumption on the normalization of p we know that

$$\left|(p_l^k)_{l,k}\right|_{\mathbb{C}^{\sum_{k\in K} d_{n,k}}} = 1.$$

A basis $\{\mathscr{R}_{B(l)}\}_{l=1,\dots,\sum_{k\in K} d_{n,k}}$ of $\mathscr{R}_K$ is constructed by setting $\mathscr{R}_{B_1} = \mathscr{R}_p$ and extending with elements of the standard basis $\{\mathscr{R}_{\mathfrak{h}_l^k}\}_{l,k}$ or a linearly independent set of rotated versions of $\mathscr{R}_p$. This basis can then be converted to an orthonormal basis of $\mathscr{R}_K$ by applying the Gram-Schmidt process to the coefficients of the basis $\{\mathscr{R}_{B(l)}\}_{l=1,\dots,\sum_{k\in K} d_{n,k}}$ with respect to the standard basis $\{\mathscr{R}_{h_l^k}\}_{l,k}$.

The approach in Remark 8.2.20 is especially useful since the operators $\mathscr{R}_p$ are similar to derivations as stated in the following.

Theorem 8.2.21 (Higher Riesz transforms and derivatives)
LET *$f \in L^2(\mathbb{R}^n)$ such that $f = \triangle^{k/2} g$ for some $g \in W^{k,2}(\mathbb{R}^n)$ and let $p \in \mathfrak{P}_k$. (Here $\triangle^{k/2} : W^{k,2(\mathbb{R}^n)} \to L^2(\mathbb{R}^n)$ is the k-th order (fractional) differential operator given by the Fourier symbol $(2\pi i)^k |\xi|^k$.)*
THEN *the partial Riesz transform $\mathscr{R}_p$ can be interpreted as an derivative in the sense that*

$$\mathscr{R}_p f = (i)^k p(\partial) g.$$

Proof.

$$\mathscr{F}\big(\mathscr{R}_p f\big)(\xi) = \frac{p(i\xi)}{|\xi|^k}\widehat{f}(\xi) = p(i\xi)\frac{(2\pi i)^k|\xi|^k}{|\xi|^k}\widehat{g}(\xi) = (i)^k \mathscr{F}\big(p(\partial)g\big)(\xi), \text{ f.a.a. } x \in \mathbb{R}^n.$$

□

We now show that steerable pyramids are based on partial higher Riesz transforms:

Remark 8.2.22 (Steerable filters and higher Riesz transforms)
To construct a steerable pyramid in 2-D the proposed filter $f : \mathbb{R}^2 \to \mathbb{R}$ is written in polar coordinates and then decomposed into a Fourier series with respect to the angular coordinate:

$$f(r,\phi) = \sum_{l=-k}^{k} a_l(r) e^{2\pi i l \phi}. \tag{8.10}$$

The minimum number of filters needed for steering is the number m of indices l for which $\exists r \in \mathbb{R}_0^+ : a_l(r) \neq 0$. As a consequence a filter is steerable iff the highest order $k < \infty$. A basis for the rotation invariant space in which the filter f lies is then given by $\{\rho_l f\}_{l=0}^{m-1}$, where ρ_l is a rotation by the angle $2\pi/m$.

Let us give an interpretation for this approach. Let $f \in L^2(\mathbb{R}^2,\mathbb{R})$ and suppose that

$$\exists \widetilde{f}:\mathbb{R}_0^+ \to \mathbb{R}: a_l(r) = \widetilde{f}(r)a_l(1). \tag{8.11}$$

Remember that $\{e^{-2\pi ik\,\cdot}, e^{2\pi ik\,\cdot}\}$ is a basis for H_k. Thus (8.10) states that

$$f \in H_f := \bigoplus_{k\in\mathbb{N}:a_k(r)\neq 0} \mathfrak{H}_k.$$

If (8.11) does not hold then $\forall l \exists f_l(r) : \mathbb{R}_0^+ \to \mathbb{R} : a_l(r) = f_l(r)a_l(1)$ and hence $f \in H_f$. Note that since f is real valued $a_l \neq 0 \Leftrightarrow a_{-l} \neq 0$. As a consequence m is the dimension of H_f. Since H_f is rotation invariant, rotating f yields an element of H_f. Let ρ_l be a rotation by $\frac{2\pi l}{m}$. To show that the m rotated versions of f yield a basis for the rotation invariant subspace $\text{span}\{\rho_l f,\ l = 0,\ldots,m\} \subset H_f$ it only remains to show that they are linearly independent - this holds true since the spaces $\mathfrak{H}_k$ are irreducible and orthogonal to each other. If (8.11) holds by Corollary 8.1.12 the filter f is the partial higher Riesz transform with respect to the polynomial $p = \sum_l a_l e^{2\pi il\,\cdot}$ of the function $\widetilde{f}$. The rotated filters are partial higher Riesz transforms of $\widetilde{f}$ with respect to rotated versions of the polynomial p.

If Equation 8.11 does not hold, then the signal f is the sum of the partial higher Riesz transforms of the function f_l with respect to the polynomials $a_l e^{-2\pi il(\,\cdot\,)}$.

8.3 Implementational aspects

In this section we discuss the implementation of higher Riesz transforms. The first step towards an implementation is of course the construction of anorthonormal bases of the spaces $\mathcal{H}_k(\mathbb{R}^n)$ and the corresponding steering matrices $\mathfrak{D}$. We discuss such constructions for the interesting cases $n = 2,3$ for which there is a proper physical interpretation for rotations. The next step is the implementation of the higher Riesz transforms via wavelet frames. This is discussed in subsection 8.3.2.

8.3.1 Explicit constructions for $L^2(\mathbb{R}^2)$ and $L^2(\mathbb{R}^3)$

In this section we will give some explicit constructions of orthonormal bases of the spaces $\mathcal{H}_k(\mathbb{R}^n)$ for $n = 2,3$.

A basis for $\mathcal{H}_k(\mathbb{R}^2)$

Theorem 8.3.1
LET $k \in \mathbb{N}$.

THEN $d_{2,k} = 2$. *A basis for $\mathcal{H}_k(\mathbb{R}^2)$ is given by*

$$\{y_1^k, y_2^k\} = \{\Re((Y(x))^k), \Im((Y(x))^k)\},$$

where $Y(x) := x_1 + ix_2$, $\forall x = (x_1,x_2) \in S^1$. Furthermore, $\langle f,g\rangle = \left\langle \frac{y_1^k}{|x|^k} f, \frac{y_1^k}{|x|^k} g\right\rangle + \left\langle \frac{y_2^k}{|x|^k} f, \frac{y_2^k}{|x|^k} g\right\rangle$.

A rotation by an angle $\phi \in [0,2\pi]$ is given by the matrix

$$\mathfrak{D}_{\rho,k} = \begin{pmatrix} y_1^k(\cos(\phi),\sin(\phi)) & -y_2^k(\cos(\phi),\sin(\phi)) \\ y_2^k(\cos(\phi),\sin(\phi)) & y_1^k(\cos(\phi),\sin(\phi)) \end{pmatrix}.$$

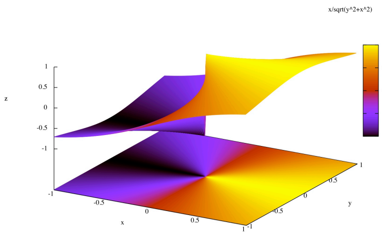
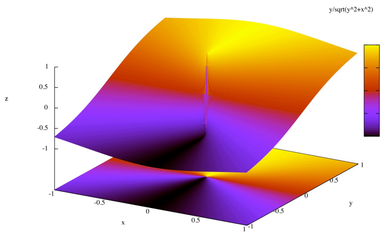
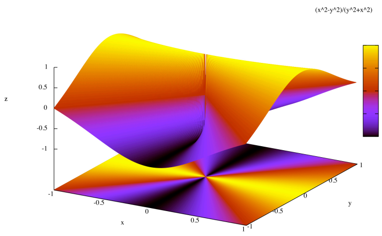
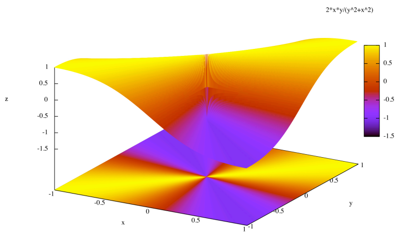
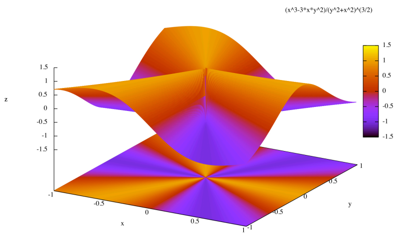
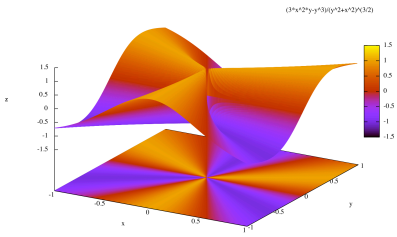
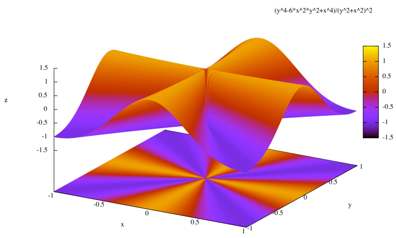
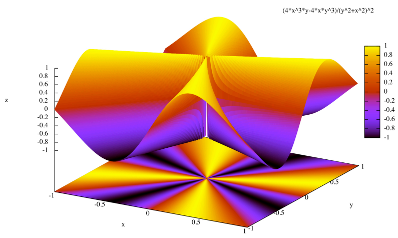
	$\boldsymbol{y_1^k}$	$\boldsymbol{y_2^k}$
$k=0$	$y_1^0(x_1,x_2)=1$	$y_2^0(x_1,x_2)=0$
$k=1$	 $y_1^1(x_1,x_2)=x_1$	 $y_2^1(x_1,x_2)=x_2$
$k=2$	 $y_1^2(x_1,x_2)=x_1^2-x_2^2$	 $y_2^2(x_1,x_2)=2x_1x_2$
$k=3$	 $y_1^3(x_1,x_2)=x_1^3-3x_1x_2^2$	 $y_2^3(x_1,x_2)=3x_2x_1^2-x_2^3$
$k=4$	 $y_1^4(x_1,x_2)=x_2^4-6x_1^2x_2^2+x_1^4$	 $y_2^4(x_1,x_2)=4x_1^3x_2-4x_2^3x_1$

Figure 8.1: The basis elements given in Example 8.3.1 of the spherical harmonics for $n=2$, $k=0,\dots,4$.

Proof. The operators are obviously self-adjoint. Hence we have to proof that

$$\left(\frac{y_1^k(x)}{|x|^k}\right)^2+\left(\frac{y_2^k(x)}{|x|^k}\right)^2=1.$$

This is implied by the addition theorem 8.1.8 or easily proven by observing that

$$\begin{aligned}\left(\frac{y_1^k(x)}{|x|^k}\right)^2+\left(\frac{y_2^k(x)}{|x|^k}\right)^2 &= |x|^{-k}\Big(\big(\Re(x_1+ix_2)^k\big)^2+\big(\Im(x_1+ix_2)^k\big)^2\Big)\\ &= |x|^{-k}(x_1+ix_2)^k\overline{(x_1+ix_2)^k} = |x|^{-k}\big\|(x_1+ix_2)^k\big\|\\ &= 1.\end{aligned}$$

A rotation of imaginary numbers by an angle ϕ is achieved by multiplication by $e^{i\phi}$. Hence

$$\mathfrak{D}_{\rho,k}(y_1^k,y_2^k)=\Big(\Re\big(e^{ik\phi}(x_1+ix_2)^k\big),\Im\big(e^{ik\phi}(x_1+ix_2)^k\big)\Big).$$

Now, the conclusion follows, since

$$\Re\big(e^{ik\phi}(x_1+ix_2)^k\big)=y_1^k\big(\cos(\phi),\sin(\phi)\,y_2^k(x_1,x_2)-y_2^k\big(\cos(\phi),\sin(\phi)\,y_2^k(x_1,x_2)\big)$$

and

$$\Im\big(e^{ik\phi}(x_1+ix_2)^k\big)=y_2^k\big(\cos(\phi),\sin(\phi)\,y_1^k(x_1,x_2)+y_1^k\big(\cos(\phi),\sin(\phi)\,y_2^k(x_1,x_2)\big).$$

□

EXAMPLE 8.3.1:

$k=0:$ $\quad y_1^0=1,\quad y_2^0=0$

$k=1:$ $\quad y_1^1=x_1,\quad y_2^1=x_2$

$k=2:$ $\quad y_1^2=x_1^2-x_2^2,\quad y_2^2=2x_1x_2$

$k=3:$ $\quad y_1^3=x_1^3-3x_2^2x_1,\quad y_2^3=3x_1^2x_2-x_2^3$

$k\in\mathbb{N}:$ Using Tchebichef polynomials we write

$$y_1^k=\frac{1}{|x|^k}\sum_{j=0}^{\lfloor\frac{k}{2}\rfloor}(-1)^j\binom{k}{2j}x_1^{2j}x_2^{k-2j},\quad y_2^k=\frac{1}{|x|^k}\sum_{j=0}^{\lceil\frac{k}{2}\rceil}(-1)^{j+1}\binom{k}{2j-1}x_2^{2j-1}x_1^{k-2j+1}$$

The resulting Higher Riesz transforms are as follows:

The identity

$k=0$ $\qquad \mathscr{R}_{y_1^0}f=f,\quad \mathscr{R}_{y_2^0}f=0,$

the Riesz transforms

$k=1$ $\qquad \widehat{\mathscr{R}_{y_1^1}f}(\xi)=\dfrac{i\xi_1}{|\xi|}\widehat{f}(\xi),\quad \widehat{\mathscr{R}_{y_2^1}f}(\xi)=\dfrac{i\xi_2}{|\xi|}\widehat{f}(\xi),$

and the higher Riesz transforms

$k\in\mathbb{N}$ $\qquad \widehat{\mathscr{R}_{y_1^k}f}(\xi)=(-i|\xi|)^{-k}y_c^k\widehat{f}(\xi),\quad \widehat{\mathscr{R}_{y_2^k}f}(\xi)=(-i|\xi|)^{-k}y_s^k\widehat{f}(\xi).$

A basis for $\mathscr{H}^k(\mathbb{R}^3)$

Normalized spherical harmonics

Definition 8.3.2
Let the $2k+1$ **associated Legendre polynomials** *of degree $k \in \mathbb{N}$ be given by*

$$P_k^m(x) := (1-x^2)^{m/2} \frac{d^{k+m}}{dx^{k+m}} (x^2-1)^k, \forall x \in \mathbb{R},$$

where $m = -k, \ldots, k$. Let $(r, \theta, \phi) \in \mathbb{R}^3$ be given in spherical coordinates – see Appendix A.1.2. Any point on the unit sphere S^2 is then determined by two coordinates $(\theta, \phi) = (1, \theta, \phi)$.
THEN *a basis for the $2k+1$ dimensional space $\mathscr{H}_k(S^2)$ of spherical harmonics of degree k is given by $\{Y_k^m\}_{m=-k,\ldots,k}$, where*

$$Y_k^m(\theta,\phi) := \sqrt{\frac{(k-m)!}{(k+m)!}} (-1)^m e^{im\phi} P_k^m(\cos(\theta)), \ \forall (\theta,\phi) \in S^2.$$

A real valued basis of spherical harmonics $\{y_k^0\} \cup \{y_{k,c}^m\}_{m=1,\ldots,k} \cup \{y_{k,s}^m\}_{m=1,\ldots,k}$ is established by setting

$$\begin{aligned}
y_k^0 &= Y_k^0(\theta,\phi) = P_k^0 \\
y_{k,c}^m &= 2^{-1/2}(Y_k^m(\theta,\phi) + (-1)^m Y_k^{-m}(\theta,\phi)) = \sqrt{2\frac{(k-m)!}{(k+m)!}} (-1)^m \cos(m\phi) P_k^m(\cos(\theta)) \\
y_{k,s}^m &= (-i)2^{-1/2}(Y_k^m(\theta,\phi) - (-1)^m Y_k^{-m}(\theta,\phi)) = \sqrt{2\frac{(k-m)!}{(k+m)!}} (-1)^m \sin(m\phi) P_k^m(\cos(\theta)),
\end{aligned}$$

$\forall m = 1, \ldots, k$, $(\theta,\phi) \in S^2$.

Wigner $\mathfrak{D}$-matrices for $\mathbb{R}^3$ In $\mathbb{R}^2$ a rotation is described by a rotation angle. In $\mathbb{R}^3$ several different approaches to describe rotations are common usage. We will consider the construction of the Wigner matrix in the case that the rotation is given in terms of Euler angles. The case that a rotation is given by an eigenvector and a rotation angle can be found in [14].

The Wigner $\mathfrak{D}$-matrix of a rotation given in Euler angles Any rotation $\rho \in SO(3)$ can be expressed as a product of three rotations which leave the z, respectively, the x-axis fixed:

$$\rho = \rho_{\alpha,z}\rho_{\beta,x}\rho_{\gamma,z} = \begin{pmatrix} \cos(\alpha) & -\sin(\alpha) & 0 \\ \sin(\alpha) & \cos(\alpha) & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(\beta) & -\sin(\beta) \\ 0 & \sin(\beta) & \cos(\beta) \end{pmatrix} \begin{pmatrix} \cos(\gamma) & -\sin(\gamma) & 0 \\ \sin(\gamma) & \cos(\gamma) & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

The irreducible representation of the rotation group is given by the Wigner-$\mathfrak{D}$ matrices [57]:

Theorem 8.3.3 (Explicit form of the Wigner-$\mathfrak{D}$ matrix)
LET *$\rho \in SO(3)$ given in Euler angels: $\rho = \rho_\alpha \rho_\beta \rho_\gamma$.*
THEN

$$Y_m^k(\rho(\theta,\phi)) = \sum_{l=-k}^{k} \left(\mathfrak{D}_{(\alpha,\beta,\gamma),k}\right)_{m,l} Y_l^k(\theta,\phi), \ \forall (\theta,\phi) \in S^2,$$

where

$$(\mathfrak{D}_{(\alpha,\beta,\gamma),k})_{m,l} = \sum_r (-1)^r \frac{c_{k,l,m} e^{im\alpha} \cos^{2k+l-m-2r}(\beta/2) \sin^{2r+m-l}(\beta/2) e^{il\gamma}}{(k-m-r)!(k+l-r)!r!(r+m-l)!}, \tag{8.12}$$

and $c_{k,l,m} = \begin{cases} (-1)^{l+m}(k-m)!(k+l)! & l,m \geq 0, \\ (k+m)!(k-l)! & l,m < 0, \\ (-1)^l (k+m)!(k+l)! & l \geq 0, m < 0, \\ (-1)^m (k-m)!(k-l)! & l < 0, m \geq 0. \end{cases}$

For $\beta = \gamma = 0$ it follows that

$$\mathfrak{D}_{(\alpha,0,0),k} = \operatorname{diag}(e^{i(m-k)\alpha}).$$

Furthermore $\mathfrak{D}_{(0,0,\gamma),k} = \mathfrak{D}_{(\gamma,0,0),k}$.

The Wigner-$\mathfrak{D}$ matrix is applied to real valued spherical harmonics in the following way:

$$\begin{aligned}
y_k^0 &= \sum_{l=1}^k \sqrt{2}\Big(y_{k,c}^l\big(\mathfrak{D}_{0,l} + (-1)^l \mathfrak{D}_{0,-l}\big) + i y_{k,s}^l \big(\mathfrak{D}_{0,l} - (-1)^l \mathfrak{D}_{0,-l}\big)\Big) + \mathfrak{D}_{0,0}\, y_k^0; \\
y_{k,c}^m &= \sum_{l=1}^k y_{k,c}^l \big(\mathfrak{D}_{m,l} + (-1)^l \mathfrak{D}_{m,-l} + (-1)^m \mathfrak{D}_{-m,l} + (-1)^{m+l} \mathfrak{D}_{-m,-l}\big) \\
&\qquad + i y_{k,s}^l \big(\mathfrak{D}_{m,l} - (-1)^l \mathfrak{D}_{m,-l} + (-1)^m \mathfrak{D}_{-m,l} - (-1)^{l+m} \mathfrak{D}_{-m,-l}\big) \\
&\quad + 2^{-1/2} y_k^0 \big(\mathfrak{D}_{m,0} + (-1)^m \mathfrak{D}_{-m,0}\big); \\
y_{k,s}^m &= \sum_{l=1}^k i y_{k,c}^l \big(\mathfrak{D}_{m,l} + (-1)^l \mathfrak{D}_{m,-l} - (-1)^m \mathfrak{D}_{-m,l} - (-1)^{l+m} \mathfrak{D}_{-m,-l}\big) \\
&\qquad - y_{k,s}^l \big(\mathfrak{D}_{m,l} - (-1)^l \mathfrak{D}_{m,-l} - (-1)^m \mathfrak{D}_{-m,l} + (-1)^{l+m} \mathfrak{D}_{-m,-l}\big) \\
&\quad + i 2^{-1/2} y_k^0 \big(\mathfrak{D}_{m,0} - (-1)^m \mathfrak{D}_{-m,0}\big).
\end{aligned}$$

8.3.2 Implementational aspects

According to Theorem 8.2.8 the higher Riesz transform of a wavelet frame is a wavelet frame for $L^2(\mathbb{R}^n)_{d_{n,k}}$. Hence we can use the same approach for the implementation of the higher Riesz transforms that we used for the Riesz transform in subsection 5.6.1 – an implementation via the isotropic wavelets constructed in Example 5.5.1. Once again the filters are implemented as perfect reconstruction filter bank in the frequency domain since they have explicit formulas and bounded support in the frequency domain. The Shannon sampling theorem ensures loss-less down-sampling after each filter step. As we constructed tight wavelet frames, the synthesis filter bank consists of the same filters as the analysis filter bank.

The Fourier multipliers of the higher Riesz transforms share a singularity at $\xi = 0$ (DC component). We deal with this singularity by subtracting the images mean value beforehand, which sets the DC component equal to zero.

The runtime of the higher Riesz wavelet decomposition is determined by that of the fast Fourier transform, namely $O(N \log N)$. The memory consumption amounts to a factor of less than $d_{n,k}$ times the image size, where $n = 2$ or $n = 3$ is the image dimension and k the degree of the higher Riesz transform.

Note that in practice the order k of the higher Riesz transforms is limited by the angular resolution of the filters - due to the discrete nature of the filters only a finite number of directions can be resolved.

8.3.3 Discussion

The present chapter gives a method to construct linear bounded operators - the higher Riesz transforms - which are invariant under translation and dilation and map wavelet frames to steerable multiwavelet frames. The invariance of the higher Riesz transforms under dilation and translation means that the dual frames are again multiwavelet frames.

The steerabilty is the manifestation of the irreducible representations of order k of the rotation group - the elements of the higher Riesz transforms are the basis elements of a linear rotation invariant space $\mathscr{R}_{\mathfrak{H}_k}$ of linear bounded operators invariant under dilation and translation on which the representation acts. Since the representation is irreducible these spaces are the minimal spaces with these properties. As a consequence any such rotation invariant space V of steerable linear bounded operators invariant under dilation and translation can be composed of the spaces $\mathscr{R}_{\mathfrak{H}_k}$. That is

$$\exists k_1, \dots, k_l \subset \mathbb{N},\ l \in \mathbb{N} :\ V = \bigoplus_{r=1}^{l} \mathscr{R}_{\mathfrak{H}_{k_r}}.$$

Furthermore, a phase direction is defined on the spaces $\mathscr{R}_{\mathfrak{H}_k}$ and on the component spaces of either only odd or only even order. Applied to the wavelet decomposition the phase direction yields a local estimate of the geometry of a signal.

8.3.4 Comparison with other approaches

In [52],[51] Michael Unser and Dimitri Van De Ville gave a definition of higher order Riesz transforms that is closely related - the elements of their higher order Riesz transforms are a basis for $\mathscr{R}_{\mathfrak{P}_k}$. Our approach is much more flexible and with the theoretical background we have set up we can give new insights into their work.

The flexibility of our work stems from the possibility to compose the elementary spaces $\mathscr{R}_{\mathfrak{H}_k}$ in an arbitrary way thus controlling symmetry and redundancy. In contrast the spaces $\mathscr{R}_{\mathfrak{P}_k}$ come in two sets of even or odd order. Within these sets the spaces of lower order are included in the spaces of higher order - it is not possible to choose one order without using all lower orders. However this is not apparent in the basis Unser et al. chose for these spaces. In contrast the basis we present is composed of sets of bases of the minimal rotation invariant subspaces and thus does allow to analyze the behavior of the signal in these subspaces as well as on the whole space. Furthermore the higher Riesz transforms we defined allow the computation of the phase direction and phase decomposition for the whole space as well as for the rotation invariant subspaces.

The theoretical background given in this chapter can explain some phenomenon remarked by Unser et al. The orthogonality of the steering matrices and their group structure remarked about in [53] is caused by the unitary representation of the rotation group.

In [53] Unser et al. propose to apply a unitary matrix pointwise to the higher order Riesz transforms. This is of course just a change of the basis of the space $\mathscr{R}_{\mathfrak{P}_k}$.

The implementation of the higher Riesz transformed wavelets is a steerable pyramid as defined by Freeman and Adelson in [18]. This is however not due to chance but due to the fact that the steerable filters designed by Freeman and Adelson are the basis of a certain rotation invariant subspace of functions. Indeed these basis elements consist of higher Riesz transforms of suitable functions.

The main difference between higher Riesz transforms and steerable pyramids is that steerable pyramids design one filter that will be a basis element of the rotation invariant space and then design a basis of rotated versions of this filter of the appropriate minimal rotation invariant space. In contrast we chose the rotation invariant spaces and then find an appropriate basis for them. Adaptivity is in our approach not part of the choice of basis - rather the phase direction of the wavelet decomposition choses the wavelet frame that is locally best adapted to the signal.

What we have found in Remark 8.2.22 is that the steerable filters of Freeman and Adelson are higher Riesz transforms. The reason why these filters work only in discrete space is the construction of the radial function $\widetilde{f}$ - they are constructed in a way to yield a filter bank. Of course it is not guaranteed that this filter corresponds to a wavelet frame of $L^2(\mathbb{R}^n)$.

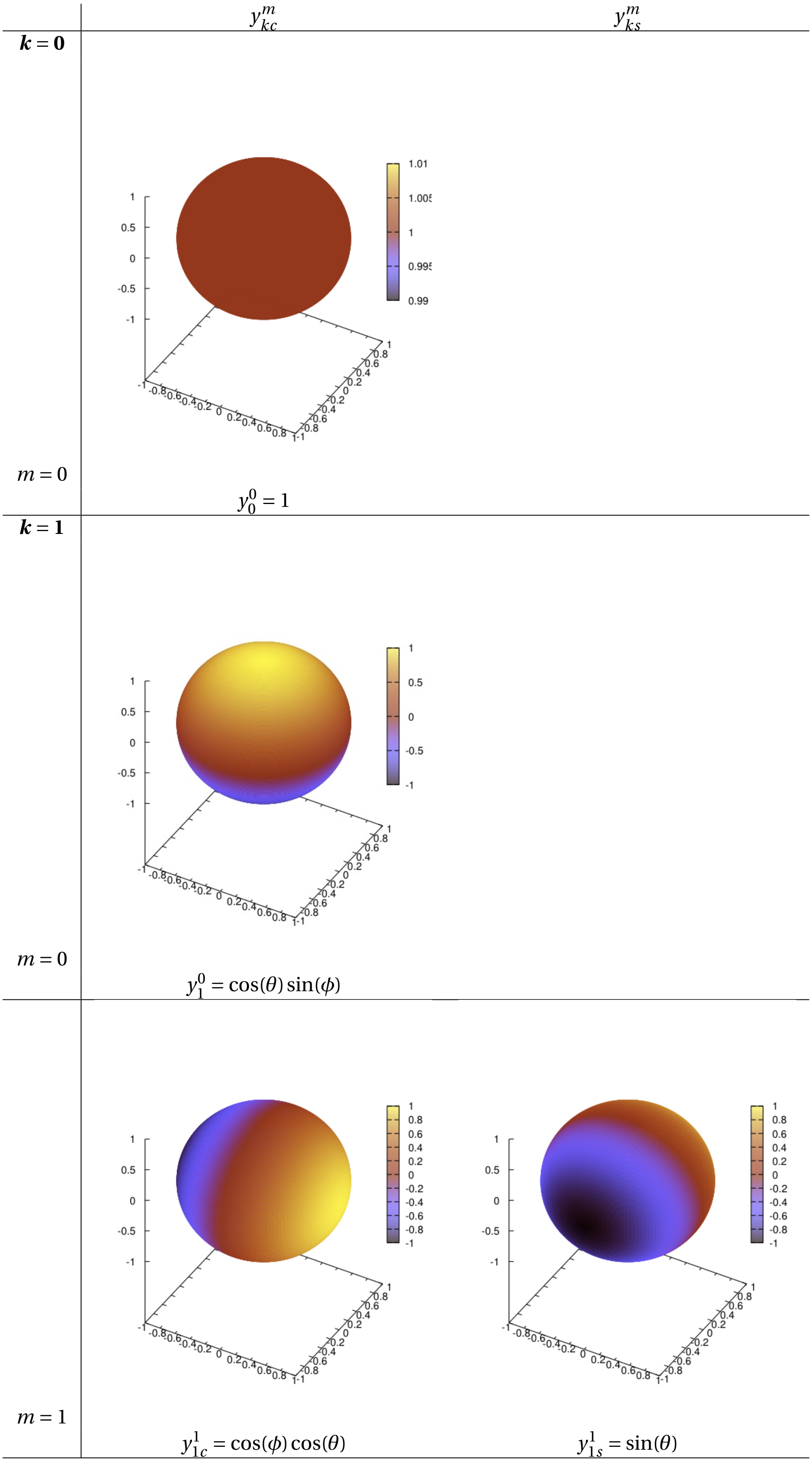


Figure 8.2: The basis elements given in Definition 8.3.2 of the spherical harmonics for $n = 3$, $k = 0, 1$.

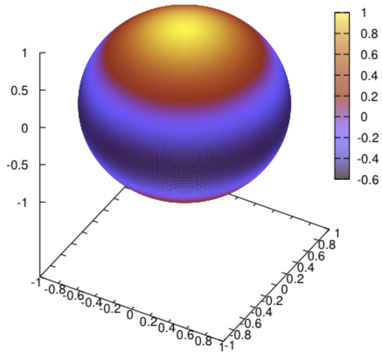
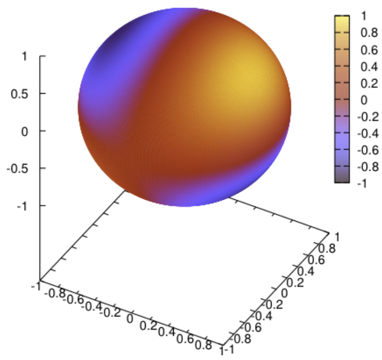
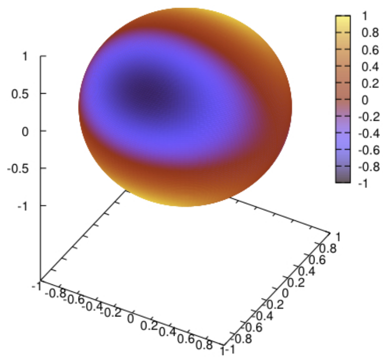
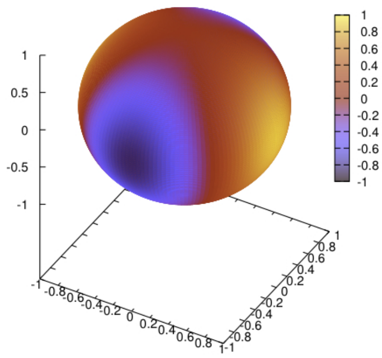
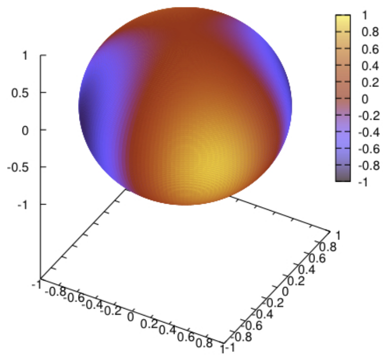
	y^m_{kc}	y^m_{ks}
$\boldsymbol{k=2}$		
$m=0$	 $y_2^0 = \frac{3\cos^2(\Theta)-1}{2}$	
$m=1$	 $y^1_{2,c} = \frac{\sqrt{6}\sin(2\phi)\cos(\theta)}{2^{3/2}}$	 $y^1_{2,s} = \frac{\sqrt{6}\sin(2\phi)\sin(\theta)}{2^{3/2}}$
$m=2$	 $y^2_{2,c} = \frac{\sqrt{6}\left(\sin(2\phi)\cos(2\theta)+\sin^2(\phi)-\cos^2(\phi)\right)}{2^{5/2}}$	 $y^2_{2,s} = \frac{\sqrt{6}\left(\sin(2\phi)\cos(2\theta)-2\cos(\phi)\sin(\phi)\right)}{2^{5/2}}$

Figure 8.3: The basis elements given in Definition 8.3.2 of the spherical harmonics for $n=3$, $k=2$.

Appendix A

The Appendix

A.1 Conventions and definitions

A.1.1 Lie groups and Lie algebras

LET L be a vector space over $\mathbb{K}$, where $\mathbb{K}$ is either $\mathbb{R}$ or $\mathbb{C}$. Furthermore let $[\,\cdot\,,\,\cdot\,]$ be a **Lie bracket**, that is, a mapping

$$[\,\cdot\,,\,\cdot\,] : L \times L \to L$$

such that for all $A, B, C \in L$ and $a, b \in \mathbb{K}$

1. $[aA + bB, C] = a[A, C] + b[B, C]$,
2. $[A, B] = -[B, A]$,
3. $[A, [B, C]] + [B, [C, A]] + [C, [A, B]] = 0$.

THEN $\big(L, [\,\cdot\,,\,\cdot\,]\big)$ is a **Lie algebra**.

LET G be a n-dimensional C^∞ manifold such that

- G is a group,
- the mapping $(\,\cdot\,,\,\cdot\,) : G \times G \to G,\ (g, h) \mapsto gh^{-1}$ is a C^∞ mapping
- there is a countable covering of G consisting of open subsets of G.

THEN G is a **Lie group**.

A.1.2 Spherical coordinates

Let $x = (x_1, x_2, x_3) = (r\cos(\phi)\cos(\theta), r\sin(\phi)\cos(\theta), r\sin(\theta)) \in \mathbb{R}^3$. Then $x = (r, \theta, \phi) \in \mathbb{R}^3$ denote the **spherical coordinates**, where $\phi \in [0, 2\pi]$, $\theta \in [0, \pi]$ and $r \in \mathbb{R}^+$.

A.1.3 Spaces of Lebesque measurable funtions

Definition A.1.1
Two functions $f, g : \mathbb{R}^n \to \mathbb{C}$ are said to be equivalent **almost everywhere**, *if they are equal except for a set of Lebesgue measure zero. (See [41] for the definition of Lebesgue measure.) Let $1 \le p < \infty$. The* **spaces of Lebesgue measurable functions** *are*

$$L^p(\mathbb{R}^n) = \left\{ f : \mathbb{R}^n \to \mathbb{C} : f \text{ is Lebesgue measurable and } \left(\int_{\mathbb{R}^n} |f(x)|^p \, dx \right)^{1/p} < \infty \right\}.$$

Here f is considered not as a single function but rather as the equivalence class of functions which coincide almost everywhere. In this context $\|f\|_{L^p} = \left(\int_{\mathbb{R}^n} |f(x)|^p \, dx \right)^{1/p}$ is a norm for $L^p(\mathbb{R}^n)$. $L^2(\mathbb{R}^n)$ is a Hilbert space with scalar product $\langle f, g \rangle := \int_{\mathbb{R}^n} f(x)\overline{g(x)} \, dx$.

A.1.4 Taylor's theorem

Definition A.1.2 (Taylor polynomial)
Let $\Omega \subseteq \mathbb{R}^n$, $a \in \Omega$ and $f \in C^k(\Omega)$. The k-th **Taylor polynomial** *is defined as*

$$T_k f(x; a) := \sum_{l=0}^{k} \frac{1}{l!} \sum_{\substack{\boldsymbol{\alpha} \in \mathbb{N}_0^n \\ |\boldsymbol{\alpha}| = l}} \partial^{\boldsymbol{\alpha}} f(a)(x-a)^{\boldsymbol{\alpha}}.$$

Theorem A.1.3 (Taylors theorem)
Let $\Omega \subseteq \mathbb{R}^n$, $a \in \Omega$ and $f \in C^k(\Omega)$.

Then

$$f(x) = T_k f(x; a) + o(|x-a|^k),$$

as $x \to a$.

That is,

$$\lim_{x \to a} \frac{f(x) - T_p f(x; a)}{|x-a|^k} = 0.$$

A.1.5 The formula of Faá di Bruno

Lemma A.1.4 (The formula of Faá di Bruno)
Let $g, h \in C^m(\mathbb{R})$. Then

$$\frac{d^m}{dt^m} g(h(t)) = \sum_{b_1, \ldots, b_m \in T_m} \frac{m!}{\prod_{k=1}^m b_k!} g^{(\sum_{k=1}^m b_k)}\big(h(t)\big) \prod_{k=1}^m \left(\frac{h^{(k)}(t)}{k!} \right)^{b_k},$$

where $T_m \subset \mathbb{N}_0^m$ is the set of sets nonnegative integers such that $m = \sum_{k=1}^m k b_k$.

(For example

$$
\begin{aligned}
&T_1 = \{\{1\}\},\\
&T_2 = \{\{1,0\},\{0,1\}\},\\
&\quad \textit{since}\; 2 = 2\cdot 1 + 0\cdot 2 = 0\cdot 1 + 1\cdot 2,\\
&T_3 = \{\{3,0,0\},\{1,1,0\},\{0,0,1\}\},\\
&\quad \textit{since}\; 3 = 3\cdot 1 = 1\cdot 1 + 1\cdot 2 = 1\cdot 3\\
&T_4 = \{\{4,0,0,0\},\{2,1,0,0\},\{1,0,1,0\},\{0,0,0,1\}\},\\
&\quad \textit{since}\; 4 = 4\cdot 1 = 2\cdot 1 + 1\cdot 2 = 1\cdot 1 + 0\cdot 2 + 1\cdot 3 = 1\cdot 4.)
\end{aligned}
$$

Proof. For a proof see [28]. □

A.1.6 Multiresolution Analysis (MRA)

Wavelet systems which constitute an orthonormal basis of $L^2(\mathbb{R}^n)$ are usually constructed starting from a multiresolution analysis.

Definition A.1.5 (Multiresolution analysis (MRA))
A **multiresolution analysis (MRA)** *is a sequence* $\{V_j\}_{j\in\mathbb{Z}}$ *of closed subspaces of* $L^2(\mathbb{R}^n)$ *such that*

1. $V_j \subset V_{j+1}$, $\forall j \in \mathbb{Z}$;
2. $\operatorname{span}\left(\bigcup_{j\in\mathbb{Z}} V_j\right) = L^2(\mathbb{R}^n)$;
3. $\bigcap_{j\in\mathbb{Z}} V_j = \{0\}$;
4. $f \in V_j \quad \Leftrightarrow \quad D_2^{-j} f \in V_0$;
5. $f \in V_0 \quad \Leftrightarrow \quad T_m f \in V_0$, $\forall m \in \mathbb{Z}^n$;
6. *there exists* $\phi \in V_0$ *such that* $\{T_m\phi\}_{m\in\mathbb{Z}^n}$ *is an orthonormal basis in* V_0. ϕ *is called a* **scaling function**.

Let $W_0 \subset V_1$ be determined by $V_1 = V_0 \oplus W_0$. A mother wavelet ψ is then a function such that $\{T_m\psi\}_{m\in\mathbb{Z}^n}$ is an orthonormal basis of W_0.

A.2 Distributions

A.2.1 Test functions

Definition A.2.1 (Test functions)

($\mathscr{S}$) *We define the Schwartz class of rapidly decreasing functions $\mathscr{S}(\mathbb{R}^n)$ as the class of functions $f \in C^\infty(\mathbb{R}^n)$, $n \in \mathbb{N}$, such that for each $j, k \in \mathbb{N}_0$ there exists a constant $C_{j,k} \in \mathbb{R}^+$ so that*
$$\rho_{j,k}(f) = \sup_{x\in\mathbb{R}^n,\,|\alpha|\le j} (1+|x|)^k |\partial^\alpha f(x)| = C_{j,k} < \infty,$$
where $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}_0^n$ is a multi-index. As usual we denote the partial derivative with respect to α by
$$\partial^\alpha f(x) = \frac{\partial^{|\alpha|}}{\partial^{\alpha_1} x_1 \cdots \partial^{\alpha_n} x_n} f(x),$$
where $|\alpha| = \sum_{l=1}^n \alpha_l$.

($\mathscr{E}$) *$\mathscr{E}(\mathbb{R}^n) := C^\infty(\mathbb{R}^n)$ where the topology is given by the set of seminorms*
$$\|\phi\|_N := \sup_{|x|\le N, |\alpha|\le N} |\partial^\alpha \phi(x)|.$$

($\mathscr{D}$) *The space of test functions $\mathscr{D}(\mathbb{R}^n)$ is given as the set of functions in $C^\infty(\mathbb{R}^n)$ that are of compact support. For a definition of a topology on $\mathscr{D}(\mathbb{R}^n)$ see for example [40] Chapter 6.2. A sequence $\{\phi_k\}_{k\in\mathbb{N}} \subset \mathscr{D}(\mathbb{R}^n)$ converges to $\phi \in \mathscr{D}(\mathbb{R}^n)$ iff there exists a compact set $K \subset \mathbb{R}^n$ such that* $\operatorname{supp}(\phi_k) \in K$, $\forall k \in \mathbb{N}$ *and* $\lim_{k\to\infty} \|\phi_k - \phi\|_n = 0$, $\forall n \in \mathbb{N}$.

Theorem A.2.2
$\mathscr{S}(\mathbb{R}^n)$ is a Fréchet space when endowed with the set of seminorms $\rho_{\alpha,k}$.
$\mathscr{S}(\mathbb{R}^n)$ is a dense subspace of L^p, $1 \le p < \infty$. $\mathscr{D}(\mathbb{R}^n)$ is a dense subspace of $\mathscr{S}(\mathbb{R}^n)$ which again is a dense subspace of $\mathscr{E}(\mathbb{R}^n)$.

Proof. The proof of the fact that $\mathscr{S}(\mathbb{R}^n)$ is a Fréchet space can be found in [40] Theorem 7.4. The space of test functions $\mathscr{D}(\mathbb{R}^n)$, is dense in $\mathscr{S}(\mathbb{R}^n)$. (See [40] 7.10) But $\mathscr{D}(\mathbb{R}^n)$ is dense in $L^p(\mathbb{R}^n)$, $1 \le p < \infty$. (See [56] Lemma V.1.9.) □

A.2.2 Distributions

Definition A.2.3 (Distributions)

($\mathscr{S}'(\mathbb{R}^n)$) *Elements of the topological dual space $\mathscr{S}'(\mathbb{R}^n)$ of $\mathscr{S}(\mathbb{R}^n)$ are called tempered distributions.*
A function $f : \mathscr{S}(\mathbb{R}^n) \to \mathbb{K}$ is an element of $\mathscr{S}'(\mathbb{R}^n)$ iff f is linear and continuous. A linear function $f : \mathscr{S}(\mathbb{R}^n) \to \mathbb{K}$ is continuous iff
$$\exists C > 0,\ j, k \in \mathbb{N}_0 : |g(\sigma)| \le C \sup_{|\alpha|\le k, x\in\mathbb{R}^n} (1+|x|)^j |D^\alpha \sigma(x)|,\ \forall \sigma \in \mathscr{S}(\mathbb{R}^n).$$

($\mathscr{E}'(\mathbb{R}^n)$) *Elements of the dual space $\mathscr{E}'(\mathbb{R}^n)$ of $\mathscr{E}(\mathbb{R}^n)$ are called distributions of compact support.*
A function $f : \mathscr{E}(\mathbb{R}^n) \to \mathbb{K}$ is element of $\mathscr{E}'(\mathbb{R}^n)$ iff f is linear and continuous. A linear function $f : \mathscr{E}(\mathbb{R}^n) \to \mathbb{K}$ is continuous iff

$$\exists N \in \mathbb{N}_0, C > 0 : |f(\eta)| \leq C \sup_{|x|\leq N, |\alpha|\leq N} |\partial^\alpha \eta(x)|, \ \forall \eta \in \mathscr{E}(\mathbb{R}^n).$$

($\mathscr{D}'(\mathbb{R}^n)$) *A function $f : \mathscr{D}(\mathbb{R}^n) \to \mathbb{K}$ is element of $\mathscr{D}'(\mathbb{R}^n)$ iff f is linear and continuous. A linear function $f : \mathscr{D}(\mathbb{R}^n) \to \mathbb{K}$ is continuous iff for every compact set $K \subset \mathbb{R}^n$*

$$\exists N \in \mathbb{N}_0, C > 0 : |f(\eta)| \leq C \sup_{x\in K, |\alpha|\leq N} |\partial^\alpha \eta(x)|, \ \forall \eta \in \mathscr{D}(\mathbb{R}^n) : \operatorname{supp}(\eta) \subset K, \qquad \text{(A.1)}$$

or equivalently iff for every convergent sequence $\{\phi_k\}_{\in\mathbb{N}} \subset \mathscr{D}(\mathbb{R}^n)$ the sequence $\{f(\phi_k)\}_{k\in\mathbb{N}} \subset \mathbb{K}$ converges.
The order N of the distribution f is the smallest number $N \in \mathbb{N}$ that will satisfy (A.1) *for all compact sets $K \subseteq \mathbb{R}^n$. If no such number exists, f is said to have infinite order.*

Definition A.2.4 (Support of a distribution and convolution with test functions)
The support of f is defined as

$$\operatorname{supp}(f) := \bigcap \{K = \overline{K} \subset \mathbb{R}^n : \operatorname{supp}(\phi) \subseteq K^c \Rightarrow \langle f, \phi \rangle = 0\}.$$

The convolution of a distribution f with a corresponding test function ϕ is defined as

$$f * \phi(y) := f(T_y \check{\phi}),$$

where $\check{\phi}(x) = \phi(-x)$, $\forall x \in (\mathbb{R}^n)$.

Theorem A.2.5 (Convolution with test functions)

(i) Let $f \in \mathscr{D}'(\mathbb{R}^n)$, $\phi \in \mathscr{D}(\mathbb{R}^n)$, then $f * \phi \in \mathscr{E}(\mathbb{R}^n)$.

(ii) Let $f \in \mathscr{E}'(\mathbb{R}^n)$, $\phi \in \mathscr{D}(\mathbb{R}^n)$, then $f * \phi \in \mathscr{D}(\mathbb{R}^n)$.

(iii) Let $f \in \mathscr{S}'(\mathbb{R}^n)$, $\phi \in \mathscr{S}(\mathbb{R}^n)$, then $f * \phi \in \mathscr{E}(\mathbb{R}^n)$, and

$$\partial^\alpha (f * \phi) = \partial^\alpha f * \phi = f * \partial^\alpha \phi.$$

Moreover $f * \phi \in \mathscr{S}'(\mathbb{R}^n)$.

(iv) Let $f, g \in \mathscr{S}'(\mathbb{R}^n) : f|_{\mathscr{D}(\mathbb{R}^n)} = g|_{\mathscr{D}(\mathbb{R}^n)} \Rightarrow f = g$

Proof. (i) For a proof see [40] Theorem 6.30 (b).

(ii) For a proof see [40] Theorem 6.35 (c).

(iii) For a proof see [40] Theorem 7.19 (a), (b).

(iv) Let $\phi \in \mathscr{S}(\mathbb{R}^n) \Rightarrow \exists \{\phi_l\}_{l\in\mathbb{N}} \subset \mathscr{D}(\mathbb{R}^n) : \phi_l \underset{\mathscr{S}(\mathbb{R}^n)}{\to} \phi$. Then

$$f(\phi) = f(\lim_{l\to\infty} \phi_l) = \lim_{l\to\infty} f(\phi_l) = \lim_{l\to\infty} g(\phi_l) = g(\phi).$$

□

A.2.3 The direct product of distributions

Definition A.2.6 (The direct product of distributions)
Let $n, m \in \mathbb{N}$, $f \in V'(\mathbb{R}^n)$, $g \in V'(\mathbb{R}^m)$, where $V = \mathscr{D}$ or $V = \mathscr{S}$. The direct product of f and g is defined as

$$(f \times g)(\phi) := f\Big(g\big(\phi(\,\cdot\,_f,\,\cdot\,_g)\big)\Big),\ \forall \phi \in V(\mathbb{R}^n \times \mathbb{R}^m).$$

Theorem A.2.7 (The direct product)
Let $l, n, m \in \mathbb{N}$, $f \in V'(\mathbb{R}^n)$, $g \in V'(\mathbb{R}^m)$, $h \in V(\mathbb{R}^l)$.

(i) The direct product given in Definition A.2.6 is well defined.

(ii) The direct product is commutative, i.e., $f \times g = g \times f$.

(iii) The direct product is associative, that is,

$$(f \times g) \times h = f \times (g \times h).$$

Proof. See [58] Theorem 5.2-1, 5.2-2, 5.3-2 and 5.3-3. □

A.2.4 Convolution of distributions

Definition A.2.8 (Convolution of distributions)
Let $f, g \in V'(\mathbb{R}^n), \phi \in V$, where $V = \mathscr{D}$ or $V = \mathscr{S}$. The convolution of f and g is defined as

$$f * g(\phi) := f \times g\big(\phi(\,\cdot\,_f + \,\cdot\,_g)\big) = f\Big(g\big(\phi(\,\cdot\,_f + \,\cdot\,_g)\big)\Big).$$

Theorem A.2.9 (Convolution of distributions)
Let $f, g \in \mathscr{D}'(\mathbb{R}^n)$ The convolution defined in Definition 1.3.8 is well defined if at least one of the distributions has compact support.

Furthermore, if f has compact support and $\exists \tilde{g} \in \mathscr{S}'(\mathbb{R}^n) : \tilde{g}|_{\mathscr{D}(\mathbb{R}^n)} = g$ then

$$\exists! T \in \mathscr{S}'(\mathbb{R}^n) :\ T|_{\mathscr{D}(\mathbb{R}^n)} = f * g. \tag{A.2}$$

Proof. We will first proof the result for $f, g \in \mathscr{D}'(\mathbb{R}^n)$. Then we will show that $|\langle f * g, \phi\rangle|$ is bounded $\forall \phi \in \mathscr{D}(\mathbb{R}^n)$ with respect to the set of seminorms on $\mathscr{S}$. From this it easily follows that there is an continuous extension of $f * g$ on $\mathscr{S}$.

Let $f, g \in \mathscr{D}'(\mathbb{R}^n)$ such that $\operatorname{supp} f$ is compact and let $h \in \mathscr{D}'(\mathbb{R}^n)$ such that there exists $\widetilde{h} \in \mathscr{S}'(\mathbb{R}^n) :\ \widetilde{h}|_{\mathscr{D}'(\mathbb{R}^n)(\mathbb{R}^n)} = h$. Furthermore let $\phi \in \mathscr{D}(\mathbb{R}^n)$, and set $\check{\phi}(x) := \phi(-x)$.

$$\begin{aligned} g * f(\phi) &= g\Big(f\big(\phi(\,\cdot\,_f + \,\cdot\,_g)\big)\Big) = g\Big(\big(f\big(\phi(\,\cdot\,_f - \,\cdot\,_g)\big)\big)^\vee\Big) \\ &= g(f * \check{\phi})^\vee. \end{aligned} \tag{A.3}$$

By Theorem A.2.5(ii), $\psi := f * \phi(\,\cdot\,) = f(T_{\cdot}\,\check{\phi}) \in \mathscr{D}(\mathbb{R}^n)$. Whence $g(f * \check{\phi})$ is well defined, implying that $g * f$ is well defined.

A similar calculation shows that $f * g(\phi) = f(g * \check{\phi})^\vee$. Furthermore, by Theorem A.2.5(i), $g * \phi \in \mathscr{E}(\mathbb{R}^n)$, and hence $f * g$ is well defined.

Moreover, $f * g = g * f$ since the direct product is commutative (Theorem A.2.7(ii)).

Let us consider $\tilde{h} \in \mathscr{S}'(\mathbb{R}^n)$. It is true that $f * h(\phi) = h(\check{\psi}) = \check{h}(\psi) = \check{\tilde{h}}(\psi)$ and hence by Definition A.2.3 we get

$$\exists a > 0,\ j, k \in \mathbb{N}_0 : |\check{\tilde{h}}(\psi)| \le a \sup_{|\alpha|\le k, x\in\mathbb{R}^n} (1+|x|)^j |D^\alpha \psi(x)|. \tag{A.4}$$

Moreover, $g \in \mathscr{D}'(\mathbb{R}^n)$ and hence for every compact set $\operatorname{supp}(\phi) \subseteq K \subset (\mathbb{R}^n)$

$$\exists N \in \mathbb{N}_0, b > 0 : |\tilde{g}(\phi)| \le b \sup_{x\in K, |\alpha|\le N} |\partial^\alpha \phi(x)|. \tag{A.5}$$

Now, by Theorem A.2.5(iii), $\partial^\beta \psi = f * \partial^\beta \check{\phi} = f(T\,.\,\partial^\beta \phi)$.

Since f has compact support, $\exists \tilde{f} \in \mathscr{E}'(\mathbb{R}^n) : \tilde{f}|_{\mathscr{D}(\mathbb{R}^n)} = f$. That is, f is linear and continuous on $\mathscr{E}(\mathbb{R}^n)$ and thus by Definition A.2.3,

$$\exists N \in \mathbb{N}_0, C > 0 : |\tilde{f}(\eta)| \le C \sup_{|x|\le N, |\alpha|\le N} |\partial^\alpha \eta(x)| \forall \eta \in \mathscr{E}(\mathbb{R}^n).$$

Hence

$$|\partial^\beta \psi(x)| = |\partial^\beta f(T_x \check{\phi})| \le C \sup_{|y|\le N, |\alpha|\le N} |\partial^{\alpha+\beta} \phi(y-x)| \tag{A.6}$$

Combined with (A.5) we obtain that for all compact sets $\operatorname{supp}(\phi) \subseteq K \subset \mathbb{R}^n$ and for all $\phi \in \mathscr{D}(\mathbb{R}^n)$

$$\exists b_{K,\phi}, C_\phi > 0, k_{K,\phi}, N_\phi \in \mathbb{N}_0 : |f * g(\phi)| = g(\psi)| \le b \sup_{|\beta|\le k, x\in K} C \sup_{|y|\le N, |\alpha|\le N} (\partial^{\alpha+\beta} \phi(y-x)).$$

Let $M \subset \mathbb{R}^n$ be compact. Since $\overline{B_N(0)}$ is compact for any $N \in \mathbb{N}_0$ we know that there exists a compact subset $K \subset \mathbb{R}^n$ and an $N \in \mathbb{N}_0$ such that $u = y - x \in M$, $\forall |y| \le N$, $x \in K$. Thus for any compact subset $M \subset \mathbb{R}^n$ and for any $\phi \in \mathscr{D}(\mathbb{R}^n)$

$$\exists c > 0,\ m \in \mathbb{N}_0 : |f * g(\phi)| \le c \sup_{u\in M, \|\alpha| \le m} |\partial^\alpha \phi(u)|.$$

Hence we proved that $f * g \in \mathscr{D}'(\mathbb{R}^n)$.

Combining (A.4) and (A.6), $\exists a, b > 0,\ j, k, N \in \mathbb{N}_0$ such that

$$|f * h(\phi)| = |\tilde{h}(\psi)| \le a \sup_{x\in\mathbb{R}^n, |\beta|\le k} (1+|x|)^j (C \sup_{|y|\le N, |\alpha|\le N} |\partial^{\alpha+\beta} \phi(y-x)|)$$

Let $u = y - x$. Then $|x| = |y - u| \le N + |u|$. Hence for all $\phi \in \mathscr{D}(\mathbb{R}^n)$

$$\exists C > 0, l, j \in \mathbb{N}_0 : |f * h(\phi)| \le C(1+|u|)^j \sup_{u\in(\mathbb{R}^n), |\alpha|\le l} |\partial^\alpha \phi(u)|. \tag{A.7}$$

Now we can show that $\exists T \in \mathscr{S}'(\mathbb{R}^n) : T|_{\mathscr{D}'(\mathbb{R}^n)} = f * h$. Let $\sigma \in \mathscr{S}(\mathbb{R}^n)$. Then there exists $\{\phi_l\}_{l\in\mathbb{N}} \subset \mathscr{D}'(\mathbb{R}^n)$ such that $\phi_l \underset{\mathscr{S}}{\to} \sigma$. Since $\{\phi_l\}$ is a Cauchy sequence, by (A.7) $\{f * h(\phi_l)\}$ is a Cauchy sequence in $\mathbb{K}$, that is, $\exists T(\sigma) := \lim_{l\to\infty} f * h(\phi_l)$. Obviously, T is a linear extension of $f * h$ on $\mathscr{S}(\mathbb{R}^n)$. Furthermore, by (A.7), $T \in \mathscr{S}'(\mathbb{R}^n)$.

It remains to show that $T \in \mathscr{S}'(\mathbb{R}^n)$ is the unique extension of $f * h$. To show this, let us assume that $\exists U \in \mathscr{S}'(\mathbb{R}^n) : U|_{\mathscr{D}(\mathbb{R}^n)} = f * h$. Let σ, ϕ_l as above. Then

$$U(\sigma) = U(\lim_{l\to\infty} (\phi_l) = \lim_{l\to\infty} f * h(\phi_l) = T(\sigma).$$

□

A.2.5 Products and distributions

Theorem A.2.10 (Products and distributions)

(i) Let $g \in \mathscr{S}'(\mathbb{R}^n)$ and let $f \in C^\infty(\mathbb{R}^n)$ such that

$$\forall \alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}_0^n \, \exists k_\alpha, C_\alpha : |(\partial^\alpha f)(x)| \le C_\alpha (1+|x|)^{k_\alpha}, \; \forall x \in \mathbb{R}^n.$$

Then the product $fg \in \mathscr{S}'(\mathbb{R}^n)$ is well defined by setting

$$fg(\phi) = g(f\phi), \; \forall \phi \in \mathscr{S}(\mathbb{R}^n).$$

(ii) Let $g \in \mathscr{D}'(\mathbb{R}^n)$. Then the product given by $fg(\phi) := g(f\phi)$ is well defined.

Proof. ad (i) The product is well defined if $f\phi \in \mathscr{S}$, $\forall \phi \in \mathscr{S}$. But this is clearly the case since any derivative of f is dominated by a polynomial.

ad (ii) The product is well defined since $\forall \phi \in \mathscr{D}(\mathbb{R}^n) \Rightarrow f\phi \in \mathscr{D}(\mathbb{R}^n)$.

□

A.2.6 Fourier transforms of distributions

Recall that in Definition 1.3.11 the Fourier transform of a function $f \in \mathscr{S}(\mathbb{R}^n)$ was defined by

$$\mathscr{F}(f)(\xi) = \widehat{f}(\xi) := \int_{\mathbb{R}^n} f(x) e^{-2\pi i x \cdot \xi} = \int_{\mathbb{R}^n} f(x) e^{-2\pi i \langle x, \xi \rangle_{\mathbb{R}^n}}.$$

The Fourier transform $\mathscr{F} : \mathscr{S}'(\mathbb{R}^n) \to \mathscr{S}'(\mathbb{R}^n)$ of a tempered distribution $f \in \mathscr{S}'(\mathbb{R}^n)$ was defined by $\widehat{f}(\phi) = f(\widehat{\phi})$, $\forall \phi \in \mathscr{S}(\mathbb{R}^n)$.

Theorem 1.3.12 stated that the Fourier transform is an isomorphism of $\mathscr{S}(\mathbb{R}^n)$. The inverse Fourier transform is given by

$$\mathscr{F}^{-1}(f)(x) := \int_{\mathbb{R}^n} f(\xi) e^{2\pi i x \cdot \xi}, \; \forall f \in \mathscr{S}'(\mathbb{R}^n), \, x \in \mathbb{R}^n.$$

The Fourier transform is a isomorphism of $\mathscr{S}'(\mathbb{R}^n)$. Its inverse on $\mathscr{S}'(\mathbb{R}^n)$ is given by

$$\mathscr{F}^{-1}(f)(\phi) = f(\mathscr{F}^{-1}\phi), \; \forall f \in \mathscr{S}'(\mathbb{R}^n), \, \phi \in \mathscr{S}(\mathbb{R}^n).$$

It follows that

$$f(\phi) = (\mathscr{F}^{-1}\mathscr{F} f)(\phi) = \mathscr{F}(f)(\mathscr{F}^{-1}\phi) = \widehat{f}(\widehat{\phi}^\vee), \; \forall f \in \mathscr{S}'(\mathbb{R}^n), \, \phi \in \mathscr{S}(\mathbb{R}^n).$$

A.2.7 A Paley-Wiener theorem

Theorem A.2.11 (Paley-Wiener)

(i) *Let $u \in \mathscr{D}'(\mathbb{R}^n)$ such that* $\operatorname{supp}(u) \subseteq \overline{B_r(0)}$, *and the order of u is N. (The order of a distribution was defined in Definition A.2.3.) Furthermore, let*

$$f(z) := u(e^{-2\pi z}),\ \forall t \in \mathbb{C}^n, \tag{A.8}$$

Then f is entire. The restriction of f to $\mathbb{R}^n$ is the Fourier transform of u and there exists a constant $0 < \gamma < \infty$ *such that*

$$|f(z)| \leq \gamma(1+|z|)^N e^{r\Im(z)},\ \forall z \in \mathbb{C}^n. \tag{A.9}$$

(ii) *Conversely, let f be an entire function in $\mathbb{C}^n$ which satisfies* (A.9) *for some* $0 < \gamma < \infty$.
Then there exists $u \in \mathscr{D}'(\mathbb{R}^n)$ with $\operatorname{supp}(u) \subseteq \overline{B_r(0)}$, *and such that* (A.8) *holds.*

Proof. The proof can be found in [40] Theorem 7.23. □

Corollary A.2.12
The Fourier transform maps $\mathscr{E}'(\mathbb{R}^n)$ to $\mathscr{E}(\mathbb{R}^n)$. The Fourier transform of a distribution $f \in \mathscr{E}'(\mathbb{R}^n)$ is given by

$$\mathscr{F}(f)(\xi) = f(e^{-2\pi i\langle\cdot,\xi\rangle}),\ \forall \xi \in \mathbb{R}^n$$

Corollary A.2.13 (Product of tempered distributions)
Let $f, g \in \mathscr{S}'(\mathbb{R}^n)$ such that f is compactly supported in the Fourier domain. Then $fg \in \mathscr{S}'(\mathbb{R}^n)$ is well defined.

Proof. Note that the statement is a simple corollary of Theorem A.2.10(i). That the necessary conditions are fulfilled is stated in Theorem A.2.11. □

A.2.8 The Fourier transform and convolution of distributions

Theorem A.2.14
Let $f, g \in \mathscr{D}'(\mathbb{R}^n)$ be such that f has compact support and $\exists \tilde{g} \in \mathscr{S}'(\mathbb{R}^n)$ such that $\tilde{g}|_{\mathscr{D}(\mathbb{R}^n)} = g$. Then $\widehat{\widetilde{f * g}} = F\widehat{\tilde{g}}$, where $F \in \mathscr{E}(\mathbb{R}^n)$ is given by $F(x) = f(e^{-2\pi i\langle\cdot,x\rangle})$ and $\widetilde{f * g} \in \mathscr{S}'(\mathbb{R}^n)$ is the unique extension of $f * g$ given in Theorem A.2.9.

Proof. Let $\phi \in \mathscr{D}(\mathbb{R}^n) \subset \mathscr{S}(\mathbb{R}^n)$. Then

$$\mathscr{F}^2(\widetilde{f * g})(\phi) = \widetilde{f * g}(\check{\phi}) = g * f(\check{\phi}) \overset{(A.3)}{=} g(f * \phi)^\vee = \tilde{g}(f * \phi)^\vee. \tag{A.10}$$

We will now look at the special case, where $[f](\phi) = \int_{\mathbb{R}^n} f(x)\phi(x)dx$ is a regular distribution and $f \in \mathscr{D}(\mathbb{R}^n)$.

It follows that

$$([f]*\phi)^\vee = \big([f](T\,.\,\check{\phi})\big)^\vee = [f](T_-\,.\,\check{\phi}) = [\check{f}](T\,.\,\phi) = \check{\widetilde{[f]}} * \check{\phi} = \widehat{\widehat{\widetilde{[f]}}} * \widehat{\widehat{\phi}} = \widehat{\widehat{f}} * \widehat{\widehat{\phi}}$$
$$= \mathscr{F}(\widehat{f}\widehat{\phi}) \in \mathscr{S}(\mathbb{R}^n).$$

Hence, using Theorem A.2.10(ii)

$$\mathscr{F}^2(\widetilde{[f]*g})(\phi) = (\widetilde{[f]*g})^\vee(\phi) = (\widetilde{[f]*g})(\check{\phi}) \overset{(A.3)}{=} g\big(([f]*\phi)^\vee\big) = \widetilde{g}(\widehat{\widehat{f}\widehat{\phi}} = \widehat{\widetilde{g}}(\widehat{f}\widehat{\phi}) = \widehat{f}\,\widehat{\widetilde{g}}(\widehat{\phi})$$
$$= \mathscr{F}(\widehat{f}\,\widehat{\widetilde{g}})(\phi),\ \forall\phi\in\mathscr{D}(\mathbb{R}^n).$$

By Theorem A.2.5(iv)

$$\forall\phi\in\mathscr{S}(\mathbb{R}^n):\ \mathscr{F}^2(\widetilde{[f]*g})(\phi) = \mathscr{F}(\widehat{f}\,\widehat{\widetilde{g}})(\phi). \tag{A.11}$$

Now let $f\in\mathscr{D}'(\mathbb{R}^n)$ have compact support and let $\phi\in\mathscr{D}(\mathbb{R}^n)$. If we define

$$[\phi]\in\mathscr{D}'(\mathbb{R}^n),\ [\phi](\eta) := \int_{\mathbb{R}^n}\phi(x)\eta(x)\,dx,\ \forall\eta\in\mathscr{D}(\mathbb{R}^n),$$

then by Theorem A.2.7(iii)

$$[\phi]*(f*g) = ([\phi]*f)*g \in \mathscr{D}'(\mathbb{R}^n).$$

By Theorem A.2.9 $\exists\widetilde{f*g}\in\mathscr{S}'(\mathbb{R}^n)$ whence $\exists\widetilde{[\phi]*(f*g)}\in\mathscr{S}'(\mathbb{R}^n)$. Furthermore, since $[\phi]$ and f have compact support, so does $f*[\phi]$. Hence, by Theorem A.2.9, there exists an extension $\widetilde{([\phi]*f)*g}\in\mathscr{S}'(\mathbb{R}^n)$.

Now Theorem A.2.5(iv) implies

$$[\phi]*(f*g) = ([\phi]*f)*g \in \mathscr{S}'(\mathbb{R}^n). \tag{A.12}$$

We first consider the left side of the equation (A.12):

(A.11) shows that

$$\mathscr{F}(\widetilde{[\phi]*(f*g)}) = \widehat{\phi}\widehat{\widetilde{f*g}} \tag{A.13}$$

Now we consider the right hand side of the equation (A.12):

We can apply (A.11) to get

$$\mathscr{F}(\widetilde{(([\phi]*f)*g}) = \widehat{f*[\phi]}\,\widehat{\widetilde{g}},. \tag{A.14}$$

which, when simplified is

$$f*[\phi](\psi) = f\Big(\int\phi(y)\psi(\,\cdot\,+y)\,dy\Big) \overset{u=\,\cdot\,+y}{=} f\Big(\int T\,.\,\phi(u)\psi(u)\,du\Big) = f\Big(\int T_u\check{\phi}(\,\cdot\,)\psi(u)\,du\Big)$$
$$= [f*\phi](\psi),\ \forall\psi\in\mathscr{D}(\mathbb{R}^n).$$

Let $\tau\in\mathscr{D}(\mathbb{R}^n):\tau(x):=1,\ \forall x\in\operatorname{supp}(f)$. Then by Theorem A.2.5(ii) $f*\phi\in\mathscr{D}(\mathbb{R}^n)$ and hence $[f*\phi]\in\mathscr{E}'(\mathbb{R}^n)$. It follows from Corollary A.2.12, that $\mathscr{F}(f*\phi)\in\mathscr{E}(\mathbb{R}^n)$. Whence

$$\widehat{f*\phi}(y) = f*\phi(e^{-2\pi i\langle\,\cdot\,,y\rangle}) = \int e^{-2\pi i\langle x,y\rangle}f(T_x\check{\phi})\,dx = f\Big(\int e^{-2\pi\langle x,y\rangle}T_x\check{\phi}\,dx\Big)$$
$$\overset{u=x-\,\cdot}{=} f\Big(\tau\int e^{-2\pi\langle(\,\cdot\,+u),y\rangle}\phi(u)\,du\Big) = f(\tau\int e^{-2\pi i\langle x,y\rangle}T_x\check{\phi}\,dx)$$
$$= f(\tau e^{-2\pi i\langle\,\cdot\,,y\rangle}\int_{\mathbb{R}^n}e^{-2\pi i\langle u,y\rangle}\phi(u)\,du) = \widehat{\phi}(y)f(e^{-2\pi i\langle y,\,\cdot\,\rangle})$$
$$= \widehat{\phi}F(y). \tag{A.15}$$

Combining (A.14), (A.15) and (A.13) we get

$$\forall \phi \in \mathscr{D}(\mathbb{R}^n): \widehat{\phi}\widehat{(\widetilde{f * g})} = \widehat{\phi} F \widehat{\widehat{g}} \in \mathscr{S}'(\mathbb{R}^n). \tag{A.16}$$

From (A.16) we now deduce our statement. Let us assume that

$$\exists \psi \in \mathscr{D}(\mathbb{R}^n): \mathscr{F}\widehat{(\widetilde{f * g})}(\psi) \neq F\widehat{\widehat{g}}(\psi).$$

Then

$$\exists a > 0 : \operatorname{supp}(\psi) \subseteq Q := \{x \in \mathbb{R}^n : |x_j| \leq a, \ \forall 1 \leq j \leq n\}.$$

Let $Q^* := \{x \in \mathbb{R}^n : |x_j| \leq b := \frac{\pi}{2a}\}$ and let $Q_\epsilon^{*c} := \{x \in \mathbb{R}^n : \nexists v \in \overline{B_\epsilon(0)}, x + v \in Q^*\}$ for some $\epsilon > 0$. Furthermore, let

$$\phi \in \mathscr{D}(\mathbb{R}^n) : \phi(x) = \begin{cases} 1, & \forall x \in Q^* \\ 0, & \forall x \in Q_\epsilon^{*c}. \end{cases}$$

Then $\widehat{\phi}(x) = \int_{Q^*} e^{-2\pi i x y} \phi(y) dy + \int_{Q*_\epsilon \setminus Q^*} e^{-2\pi i x y} \phi(y) dy =: I_1 + I_2$.

$$I_1 = \prod_{l=1}^{n} \int_{-b}^{b} e^{-2\pi i x_l t} dt = \prod_{l=1}^{n} \frac{2\sin(x_l b)}{x_l}$$

Hence, using Taylor series

$$|I_1| \geq (2b)^n \prod_{l=1}^{n} (1 - \frac{|x_l b|^2}{3!}) \geq (2b)^n (1 - \frac{\pi^2}{2^2 3!})^n > (2b)^n 2^{-n}.$$

Now we can choose $\epsilon > 0 : |I_2| < (2b)^n 2^{-n-1}$ to conclude

$$\widehat{\phi}(x) \neq 0, \ \forall x \in Q.$$

It follows that $\exists \Psi \in \mathscr{D}(\mathbb{R}^n) : \psi = \widehat{\phi}\Psi$ and hence

$$\mathscr{F}(\widetilde{f * g})(\psi) = \mathscr{F}(\widetilde{f * g})(\widehat{\phi}\Psi) = \left(\widehat{\phi}\mathscr{F}(\widetilde{f * g})\right)(\Psi) \overset{(A.16)}{=} (\widehat{\phi} F \widehat{\widehat{g}})(\Psi) = (F\widehat{\widehat{g}})(\psi).$$

which contraicts our assumption. Hence $\mathscr{F}\widehat{(\widetilde{f * g})}(\psi) = F\widehat{\widehat{g}}(\psi)$ holds $\forall \psi \in \mathscr{D}(\mathbb{R}^n)$ and hence by Theorem A.2.5(iv) $\forall \psi \in \mathscr{S}(\mathbb{R}^n)$. □

List of Figures

Bibliography

[1] Edward Adelson. Checkershadow illusion. 106, 108

[2] Akram Aldroubi, Carlos Cabrelli, and Ursula Molter. Wavelets on irregular grids with arbitrary dilation matrices, and frame atoms for $L^2(\mathbb{R}^d)$. *Applied and Computational Harmonic Analysis*, 17(2):119–140, 2005. 91

[3] J.P. Antoine, P. Carrette, R. Murenzi, and B. Piette. Image analysis with two-dimensional continuous wavelet transform. *Signal Processing*, 31(3):241–272, 1993. 127

[4] R. Azenscott, B. Bodmann, and M. Papadakis. Steerlets: A novel approach to rigid-motion covariant multiscale transforms. In *Proceedings of the SPIE: Wavelets XIII*, 2009. 2

[5] E. Bedrosian. A product theorem for hilbert transforms. *Proceedings of the IEEE*, 51:868–869, 1963. 36

[6] Bracewell. *The Fourier Transform and its Applications*. McGraw-Hill, 1965. 10

[7] F. Brackx, R. Delanghe, and F. Sommen. *Clifford Analysis*. Pitman Books Limited, 1982. 46, 47

[8] Thomas Bülow. *Hypercomplex Spectral Signal Representations for Image Processing and Analysis*. PhD thesis, University of Kiel, 1999. 76

[9] Wai Lam Chan, Hyekho Choi, and Richard G. Baraniuk. Coherent image processing using quaternion wavelets. In *Proceedings SPIE 5914*, 2005. 76

[10] Ole Christensen. *An introduction to frames and Riesz bases*. ANHA. Birkhäuser, 2003. 13, 80, 87

[11] A.L. Chystyakov. On uncertainty relations for vector-valued operators. *Theor. Math. Phys.*, 27:380–382, 1976. 129

[12] S. Dahlke and Peter Maaß. The affine uncertainty principle in one and two dimensions. *Computers Math. Applic.*, 30(3-6):293–305, 1995. 113

[13] R. Delanghe. Clifford analysis: History and perspective. *Computational Methods and Function Theory*, 1(1):107–153, 2001. 58

[14] A. Erdélyi, editor. *Bateman Manuscript Project*, volume 2. McGraw-Hill Book Company, 1953. 138, 140, 159

[15] Michael Felsberg. *Low-Level Image Processing with the Structure Multivector*. PhD thesis, Christian-Albrechts-Universität Kiel, 2002. 2, 68

[16] Michael Felsberg. Optical flow estimation from monogenic phase. In *Complex Motion*, volume 3417/2007, 2007. 106, 108

[17] Gerald B. Folland and Alladi Sitaram. The uncertainty principle: A mathematical survey. *The Journal of Fourier Analysis and Applications*, 3(3):207–138, 1997. 115, 116, 123, 125, 129

[18] William T. Freeman and Edward H. Adelson. The design and use of steerable filters. *IEEE Trans. Patt. Anal. and Machine Intell.*, 13(9), Sept. 1991. iii, 1, 3, 6, 7, 137, 154, 161

[19] John Gilbert and Margaret Murray. *Clifford algebras and Dirac operators in harmonic analysis*, volume 26 of *Cambridge studies in advanced mathematics*. Cambridge University Press, 1991. 41, 43, 58, 73, 74

[20] Karlheinz Gröchenig. *Foundations of Time-Frequency Analysis*. ANHA. Birkhäuser, 2001. 10, 11, 90, 113

[21] Klaus Gürlebeck, Klaus Habetha, and Wolfgang Sprössig. *Funktionentheorie in der Ebene und im Raum*. Birkhäuser, 2006. 41, 58

[22] Stefan Held. Uncertainty relations for sets of self adjoint operators. *Proceedings in Applied Mathematics and Mechanics*, Volume 11, Issue 1:849–850, 2011. 111

[23] Stefan Held and Martin Storath. Monogenic wavelet frames. In *Wavelets XIII, Proceedings of Spie Volume 7446*, 2009. 2, 4, 5, 65, 77

[24] Stefan. Held, Martin Storath, Brigitte Forster, and Peter Massopust. Steerable wavelet frames based on the riesz transform. *IEEE Transactions on Image Processing*, 19(3):653–667, 2010. 2, 4, 5, 65, 77, 107

[25] M. Hohlschneider. *Wavelets*. Oxford University Press, 1995. 28

[26] Lars Hörmander. *The analysis of linear partial differential operators I*. Springer, 2nd edition, 1990. 114

[27] Abdul Jerri. *The Gibbs Phenomenon in Fourier Analysis, Splines and Wavelet Approximations*. Kluwer Academic Publishers, 1998. 131, 132

[28] Warren Johnson. The curious history of Faà di bruno's formula. *The American Mathematical Monthly*, March 2002. 167

[29] Tosio Kato. *Perturbation theory for linear operators*. Springer, corrected version of the 2nd edition edition, 1995. 85

[30] F.W. King. *Hilbert Transforms*, volume 1 and 2. Cambridge University Press, 2009. 17, 27

[31] Takahide Kurokawa. On the closure of the Lizorkin space in spaces of Beppo Levi type. *Studia mathematica*, 150(2):99–120, 2002. 28

[32] Marius Mitrea. *Clifford Wavelets, Singular Integrals and Hardy spaces*, volume 1575 of *Lecture Notes in Mathematics*. Springer, 1994. 46, 49, 50, 52

[33] Claus Müller. *Spherical Harmonics*, volume 17 of *Lecture Notes in Mathematics*. Springer, 1966. 138

[34] Sofia Olhede and Georgios Metikas. The hyperanalytic wavelet transform. Technical report, Department of Mathematics, Imperial CollegeLondon, 2006. 76

[35] Alan V. Oppenheim and Jae S. Lim. The importance of phase in signals. *Proceedings of the IEEE*, 69(5), 1981. 16

[36] Wayne Rasband. ImageJ. 103

[37] Michael Reed and Barry Simon. *Methods of modern mathematical physics I: Functional Analysis*. Academic Press, 1980. 8

[38] Juan R. Romero, Simon K. Alexander, Sikha Baid, and Manos Papadakis. The geometry and the analytic properties of isotropic multiresolution analysis. *Advances in Computational Mathematics*, 2009. 2, 101, 102

[39] Amos Ron and Zuowei Shen. Affine systems in $L_2(\mathbb{R}^d)$: The analysis of the analysis operator. *J. Functional Anal*, 148:408–447, 1996. 98, 99

[40] Walter Rudin. *Functional analysis*. McGraw-Hill, 2nd edition, 1991. 9, 168, 169, 173

[41] Walter Rudin. *Real and Complex Analysis dt*. Oldenburg Wissenschaftsverlag, 3rd ed. edition, 1999. 8, 25, 71, 94, 150, 166

[42] Veronika Schlager, Stefan Schausberger, David Stifter, and Bettina Heise. Coherence probe microscopy imaging and analysis for fiber-reinforced polymers. In *Proceedings of the 17th Scandinavian conference on Image Analysis*, 2011. 2

[43] Ivan W. Selesnick, Richard G. Baraniuk, and Nick G. Kingsbury. The dual-tree complex wavelet transform. *IEEE Signal Processing Magazine*, November 2005. 1

[44] Hong-Tae Shim, Hans Volkmer, and Gilbert G. Walter. Gibbs phenomenon in higher dimensions. *Journal of Approximation Theory*, 145:20–32, 2007. 132, 134, 135

[45] Elias Stein. *Singular Integrals and Differentiability Properties of Functions*. Princeton University Press, 1970. 2, 21, 22, 24, 72, 138, 139, 141, 142, 143

[46] Elias Stein and Guido Weiss. *Introduction to Fourier analysis on Euclidean spaces*. Princeton University Press, 1971. 2, 139, 141

[47] Martin Storath. Wavelets und das monogene Signal. Master's thesis, Technische Universität München, 2008. 5, 77, 103, 106

[48] Martin Storath. The monogenic curvelet transform. In *17th IEEE International Conference on Image Processing (ICIP)*, 2010. 2

[49] Martin Storath. Directional multiscale amplitude and phase decomposition by the monogenic curvelet transform. *SIAM J. Imaging Sci.*, 4(1):57–78, 2011. 2

[50] Martin Storath and Stefan Held. Monogenic Wavelet Toolbox. 107

[51] Michael Unser and D. Van De Ville. Wavelet steerability and the higher order Riesz transform. *IEEE Transactions on Image Processing*, 2010. 2, 161

[52] Michael Unser and Dimitri Van de Ville. Higher-order Riesz tansforms and steerable wavelet frames. *IEEE Transactions on Image Processing*, 48(34), 2009. 2, 161

[53] Michael Unser, Dimitri Van De Ville, and Nicolas Chenouard. Steerable pyramids and tight wavelet frames in $L^2(\mathbb{R}^d)$. *IEEE Transactions on Image Processing*, 20(10):2705–2721, 2011. 2, 161

[54] M. Venouziou and H. Zhang. Characterization of the Hilbert transform by the Bedrosian theorem. *J. Math. Anal. Appl.*, 338:1477–1481, 2008. 39

[55] Guido Weiss and Eugenio Hernandez. *A first course on Wavelets*. CRC Press, 1996. 96

[56] Dirk Werner. *Funktionalanalysis*. Springer, 4 edition, 2002. 48, 49, 52, 57, 81, 168

[57] Eugen Wigner. *Gruppentheorie und ihre Anwendung auf die Quantenmechanik der Atomspektren*. Vieweg, 1931. 138, 144, 159

[58] A. Zemanian. *Distribution theory and transform analysis*. Dover Publications, INC., 1987. 9, 170

[59] Javier Duoandikoetxea Zuazo. *Fourier Analysis*. AMS, 2001. 15

Index