Multi-Scale Riemann-Finsler Geometry
Applications to Diffusion Tensor Imaging
and High Angular Resolution Diffusion Imaging
Colophon

Cover picture by Mandi Astola, cover design by Paul Verspaget.

This thesis was typeset by the author using \LaTeX{}.

The Netherlands Organisation for Scientific Research (NWO) is gratefully acknowledged for financial support.

This work was part of NWO vici project “The Problem of Scale in Biomedical Image Analysis”, project number 639.023.403.

Printed by PrintService TU/e, Eindhoven, the Netherlands.

A catalogue record is available from the Eindhoven University of Technology Library.


© 2010 L.J. Astola Eindhoven, The Netherlands, unless stated otherwise on chapter front pages, all rights are reserved. No part of this publication may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording, or any information storage and retrieval system, without permission in writing from the copyright owner.
Multi-Scale Riemann-Finsler Geometry
Applications to Diffusion Tensor Imaging
and High Angular Resolution Diffusion Imaging

PROEFSCHRIFT

ter verkrijging van de graad van doctor aan de Technische Universiteit Eindhoven, op gezag van de rector magnificus, prof.dr.ir. C.J. van Duijn, voor een commissie aangewezen door het College voor Promoties in het openbaar te verdedigen op woensdag 27 januari 2010 om 16.00 uur

doort

Laura Johanna Astola

deboren te Turku, Finland
Dit proefschrift is goedgekeurd door de promotor:

prof.dr. L.M.J. Florack
Contents

Colophon ii

Contents v

1 Introduction 1

2 Basics of Riemannian Geometry 5
  2.1 Connection ........................................... 6
  2.2 Curvature ............................................. 9
  2.3 Geodesics With Fixed Initial Point and
      Direction ............................................. 15
  2.4 Geodesic Deviation .................................... 16
     2.4.1 Exponential Mapping ............................. 16
     2.4.2 Geodesic Deviation .............................. 21
  2.5 Geodesics With Fixed End Points ..................... 24
     2.5.1 The Hamilton-Jacobi Equation for a Non-Homogeneous
            Integrand ......................................... 27
     2.5.2 The Hamilton-Jacobi Equation for a Homogeneous
            Integrand ......................................... 29
     2.5.3 Asymptotically Homogeneous Case Study ........... 30

3 Applications in DTI 33
  3.1 Introduction .......................................... 34
  3.2 Streamline Tracking of DTI Fibers .................... 38
  3.3 Dijkstra's Shortest Paths ............................. 41
  3.4 Quality Measures for DTI Fibers ..................... 41
     3.4.1 Connectivity Strength of Pathways ................. 41
     3.4.2 Ricci Curvature .................................. 43
     3.4.3 Stickiness ....................................... 45
     3.4.4 Inhomogeneity Detection ......................... 46
  3.5 Ricci Scalar .......................................... 47
     3.5.1 Experimental Results ............................ 49

4 Basics of Finsler Geometry 53
  4.1 Introduction .......................................... 54
  4.2 Homogeneity of Finsler Norm ........................ 55
  4.3 Strong convexity of Finsler Norm ..................... 55
     4.3.1 Strong convexity criterion ....................... 55
Introduction
This thesis studies some concepts and tools of differential geometry, namely of Riemann and Finsler geometry and applies these to medical image analysis.

The medical images considered in this thesis are diffusion tensor images (DTI) and high angular resolution diffusion images (HARDI) of the brain tissue. These are examples of medical images acquired in a magnetic resonance imaging (MRI) scanner, and are so-called non-invasive imaging techniques meaning that they are relatively harmless compared to techniques that use x-rays or contrast agents. They are designed to capture the average profile of the Brownian motion of water molecules in tissue.

Just like a two dimensional image consists of an array of pixels, a DTI or a HARDI image consists of three dimensional array of voxels. Whereas in a two dimensional digital photograph each pixel contains information on color and brightness, each voxel in DTI/HARDI image contains information on the water diffusion profile. For example in the ventricles, which are filled with cerebrospinal fluid to cushion and support the brain, the diffusion profiles look like a sphere, because water molecules can freely diffuse in all directions. On the other hand, for example in the optic radiation, which is a bundle of nerves connecting the eyes to the visual cortex, diffusion profiles look typically like elongated ellipses, because the molecules are forced to diffuse along the nerve bundle. So by looking at how an ensemble of water molecules behaves, we may try to recover the surrounding network of nerve bundles.

The analysis of DTI/HARDI-images has many useful applications. It can help in understanding how different parts of the brain are connected to each other. This is crucial in planning a neurosurgery. The surgeon must know the locations of vital nerve bundles to be able to avoid these during an operation. It can also help in diagnosis by indicating abnormalities that are related to stroke, schizophrenia, Alzheimer’s disease etc.

Most of the tools that we use in probing the data are standard equipment in Riemann and Finsler geometry. Riemann geometry is more familiar, e.g. from its applications to Einstein’s general relativity theory. Finsler geometry can be seen as a more general geometry that covers also the standard Riemann geometry. These are both well established subjects in differential geometry and are used in local as well as global analysis of manifolds. The application of Riemann-Finsler geometry to tensor valued medical image data is the main topic and novelty of this thesis.
Since MRI images contain noise, induced by subject movements or disturbing currents, it is important to smoothen the images and suppress the noise. Therefore we also derive regularization algorithms suitable for this special kind of data.

The main application considered in this thesis is in medical imaging. However, the framework presented in this thesis is by no means restricted to medical images.

In Chapter 2 the relevant basic definitions in Riemann geometry are introduced. Some concepts are studied in detail, aiming at intuitive and visual explanations of the geometric ideas involved. In Chapter 3 algorithms and measures to analyze diffusion tensor images using Riemann geometry are derived, implemented and applied. In Chapter 4 the fundamental concepts in Finsler geometry are defined and a new formulation for the convexity criterion is derived. In Chapter 5 novel applications for HARDI using Finsler geometry are proposed. Explicit algorithms are provided and experimental results are shown. In Chapter 6 new examples of Gaussian scale spaces for DTI- and HARDI-images are introduced.
Basics of Riemannian Geometry

“...and to do mathematics you have to feel comfortable and confident.”
Mary E. Rudin
2.1 Connection

In this section we introduce the fundamental concepts of Riemannian geometry, that we use for applications in later chapters.

**Definition 2.1.** Let $M$ be a differentiable manifold (for details see Definition 8.1) and let $TM$ be the set of all smooth vector fields on $M$. A connection, or a covariant derivative on $M$ is a mapping

$$\nabla : TM \times TM \rightarrow TM$$

$$(X,Y) \mapsto \nabla_X Y,$$  \hspace{1cm} (2.1)

that satisfies the following

1. $\nabla_{fX_1 + gX_2}Y = f\nabla_{X_1}Y + g\nabla_{X_2}Y,$
2. $\nabla_X(aY_1 + bY_2) = a\nabla_X Y_1 + b\nabla_X Y_2$, for $a, b \in \mathbb{R},$
3. $\nabla_X(fY) = f\nabla_X Y + (Xf)Y,$

where $f, g$ are $C^{\infty}$ real valued functions on $M$.

Let $(E_1, \ldots, E_n)$ be a local frame, i.e. a set of smoothly varying basis vectors, for the tangent space $TM$ on an open subset $U \subset M$. We take as the local frame, the coordinate frame $(\frac{\partial}{\partial x^i})$, (which we will sometimes write shortly as $\partial_i = \frac{\partial}{\partial x^i}$).

The covariant derivative of a local frame vector can be written as a linear combination of local basis vectors$^1$:

$$\nabla_{E_j}E_i = \Gamma^k_{ji}E_k.$$  \hspace{1cm} (2.2)

In coordinate frame, this is denoted as

$$\nabla_{\partial_j}\partial_i = \Gamma^k_{ji}\partial_k.$$  \hspace{1cm} (2.3)

$^1$Sometimes this is written $\nabla_{E_j}E_i = \Gamma^k_{ij}E_k$, but in the symmetric case used later, it makes no difference.
2.1. Connection

Here the Christoffel symbols $\Gamma^k_{ij}$ are formal coefficients of the derivative expressed as linear combination of local coordinate vectors. The covariant derivative of a general vector $Y$ with respect to a vector $X$ is

$$\nabla_X Y = X^i \nabla_{\partial_i} (Y^j \partial_j) = X^i (\partial_i Y^k + Y^j \Gamma^k_{ij}) \partial_k,$$

(2.4)

where the following Einstein summation convention was used (and will be used throughout this thesis)

$$a^i b^i := \sum_i a^i b^i. \quad (2.5)$$

In the following, the conditions that provide a manifold with a unique metric are introduced.

**Definition 2.2.** The Lie bracket $[X, Y]$ of two vector fields $X$ and $Y$, is the vector field corresponding to the following derivation

$$[X, Y] f = X(Y f) - Y(X f) = \left( X^j \partial_j Y^i - Y^j \partial_j X^i \right) \partial_i f. \quad (2.6)$$

**Definition 2.3.** The torsion tensor of a connection is defined as

$$\tau(X, Y) = \nabla_X Y - \nabla_Y X - [X, Y]. \quad (2.7)$$

Componentwise this is

$$\tau(X, Y) = \nabla_X Y - \nabla_Y X - [X, Y]$$

$$= \nabla_{X^i \partial_i} Y^j \partial_j - \nabla_{Y^i \partial_i} X^j \partial_j - \left( X^i \partial_i (Y^j) \partial_j - Y^i \partial_i (X^j) \partial_j \right)$$

$$= X^i \nabla_{\partial_i} (Y^j \partial_j) - Y^j \nabla_{\partial_j} (X^i \partial_i) - X^i \partial_i (Y^j \partial_j) + Y^i \partial_i (X^j \partial_j)$$

$$= X^i \partial_i Y^j \partial_j + X^i Y^j \Gamma^k_{ij} \partial_k - Y^j \partial_j X^i \partial_i + Y^j X^i \Gamma^k_{ji} \partial_k - X^i \partial_i Y^j \partial_j + Y^i \partial_i X^j \partial_j$$

$$= X^i Y^j \Gamma^k_{ij} \partial_k - Y^j X^i \Gamma^k_{ji} \partial_k$$

$$= X^i Y^j (\Gamma^k_{ij} - \Gamma^k_{ji}) \partial_k. \quad (2.8)$$

**Definition 2.4.** A linear connection is said to be symmetric if the torsion vanishes i.e. if

$$\nabla_X Y - \nabla_Y X - [X, Y] = 0,$$

for every $X, Y \in T(M)$. 
From the componentwise presentation of the torsion tensor (2.8), it can be seen that a linear connection is symmetric if and only if
\[ \Gamma^k_{ij} = \Gamma^k_{ji}. \]

**Definition 2.5.** Let \( M \) be a differentiable manifold with Riemannian metric \( \langle \cdot, \cdot \rangle \) and a connection \( \nabla \). A connection \( \nabla \) is said to be compatible with the metric \( \langle \cdot, \cdot \rangle \), if for all vector fields \( X, Y, Z \) it satisfies the product rule
\[ \nabla_X \langle Y, Z \rangle = \langle \nabla_X Y, Z \rangle + \langle Y, \nabla_X Z \rangle. \quad (2.9) \]

**Theorem 2.6 (The Fundamental Lemma of Riemannian Geometry).** Let \( (M, g) \) be a Riemannian manifold. There exists a unique linear connection \( \nabla \) on \( M \) that is symmetric and compatible with \( g \).

Following [19], we show the uniqueness from compatibility and symmetry. Since \( X, Y, Z \) in definition 2.5 are arbitrary, we can write
\[
\begin{align*}
\nabla_X \langle Y, Z \rangle &= \langle \nabla_X Y, Z \rangle + \langle Y, \nabla_X Z \rangle, \\
\nabla_Y \langle Z, X \rangle &= \langle \nabla_Y Z, X \rangle + \langle Z, \nabla_Y X \rangle, \\
\nabla_Z \langle X, Y \rangle &= \langle \nabla_Z X, Y \rangle + \langle X, \nabla_Z Y \rangle.
\end{align*}
\]

Using the symmetry condition (definition 2.4) on the last terms of each equation, and bi-linearity of the inner product
\[
\begin{align*}
\nabla_X \langle Y, Z \rangle &= \langle \nabla_X Y, Z \rangle + \langle Y, \nabla_X Z \rangle + \langle Y, [X, Z] \rangle, \\
\nabla_Y \langle Z, X \rangle &= \langle \nabla_Y Z, X \rangle + \langle Z, \nabla_Y X \rangle + \langle Z, [Y, X] \rangle, \\
\nabla_Z \langle X, Y \rangle &= \langle \nabla_Z X, Y \rangle + \langle X, \nabla_Z Y \rangle + \langle X, [Z, Y] \rangle.
\end{align*}
\]

Adding (2.11) and (2.12), and subtracting (2.13) from this, one has
\[ \langle \nabla_X Y, Z \rangle = \frac{1}{2} \left( X \langle Y, Z \rangle + Y \langle X, Z \rangle - Z \langle X, Y \rangle \right) - (\langle Y, [X, Z] \rangle - \langle Z, [Y, X] \rangle + \langle X, [Z, Y] \rangle). \quad (2.14) \]

Therefore \( \nabla \) is uniquely determined by \( \langle \cdot, \cdot \rangle \). On the other hand, if \( \nabla \) is defined by 2.14, it gives a symmetric and compatible connection proving the existence.

On a Riemannian manifold, with the Riemannian metric tensor \( g_{ij} \), the Christoffel symbols \( \Gamma^k_{ij} \) can be calculated from the identity (2.14).
\[ \langle \nabla_{\partial_i} \partial_j, \partial_l \rangle = \frac{1}{2} \left( \partial_i \langle \partial_j, \partial_l \rangle + \partial_j \langle \partial_l, \partial_i \rangle - \partial_l \langle \partial_i, \partial_j \rangle \right). \quad (2.15) \]
Since \( g_{ml} = \langle \partial_m, \partial_l \rangle \), the previous equation (2.15) becomes
\[
\Gamma^k_{ij} = \frac{1}{2} g^{kl} \left\{ \frac{\partial g_{jl}}{\partial x^i} + \frac{\partial g_{li}}{\partial x^j} - \frac{\partial g_{ij}}{\partial x^l} \right\}.
\tag{2.16}
\]
This symmetric and linear connection is also called the Levi-Civita connection.

### 2.2 Curvature

**Definition 2.7.** Let \( TM \) be the set of all \((C^\infty)\) vector fields on \( M \) and let \( X, Y, Z \in TM \). The curvature tensor \( R \) on a Riemannian manifold is a map
\[
R(X, Y) : TM \to TM
\]
\[
R(X, Y)Z = \nabla_Y \nabla_X Z - \nabla_X \nabla_Y Z - \nabla_{[X,Y]}Z.
\tag{2.17}
\]

**Definition 2.8.** The components of the curvature tensor in local coordinates are defined as
\[
R^l_{ijk} = \frac{\partial}{\partial x^l} \circ R \left( \frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j} \right) \frac{\partial}{\partial x^k},
\tag{2.18}
\]
and the components of Riemann tensor as
\[
R_{ijks} := g_{ms} R^m_{ijk} = \langle R \left( \frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j} \right) \frac{\partial}{\partial x^k}, \frac{\partial}{\partial x^s} \rangle.
\tag{2.19}
\]

Using Christoffel symbols (2.19) is
\[
g_{ms} R^m_{ijk} = g_{ms} \left( \Gamma^l_{ik} \Gamma^m_{jl} + \frac{\partial}{\partial x^j} \Gamma^m_{ik} - \Gamma^l_{jk} \Gamma^m_{il} - \frac{\partial}{\partial x^i} \Gamma^m_{jk} \right).
\tag{2.20}
\]

The curvature tensor measures non-commutativity of the covariant derivative. This can also be interpreted as measuring how much \( M \) deviates from being Euclidean.

Let \( n \) be the dimension of the Riemann manifold. In principle there are \( n^4 \) permutations of the indices of Riemann tensor, but not all of them are independent, since the Riemann tensor has the following symmetries
\[
R_{ijks} = -R_{jiks}, \quad R_{ijks} = -R_{ijsk},
\tag{2.21}
\]
\[ R_{ijks} = R_{kxis}, \quad (2.22) \]

and

\[ R_{ijks} + R_{jks} + R_{kis} = 0. \quad (2.23) \]

The symmetries (2.21) indicate a system $\boxed{i \quad j \quad k \quad s}$, where the circled indices can change place within each rectangle without introducing an independent curvature scalar. If the circled indices in one square are identical the curvature is zero, so the only possibilities are the pairs $(i, j), (k, s), (i \neq j, k \neq s)$ and the number $N$ of such combinations is

\[ N = m^2 = \binom{n}{2} \binom{n}{2} = \frac{1}{4} n^2 (n - 1)^2, \quad (2.24) \]

where $m = \binom{n}{2}$. The symmetry (2.22) states that changing the mutual positions of the rectangles does not affect the value, i.e. that $\boxed{i \quad j \quad k \quad s} \sim \boxed{k \quad s \quad i \quad j}$. Therefore half of the permutations counted in $N$, such that $\boxed{i \quad j \quad k \quad s} \neq \boxed{k \quad s \quad i \quad j}$ is redundant, and the number of these redundant elements is:

\[ \binom{m}{2} = \frac{1}{2} m(m - 1) = \frac{1}{8} n(n - 1)(n^2 - n - 2). \quad (2.25) \]

It can be easily seen that the symmetry (2.23) is relevant only in case all indices are distinct. Taking symmetries (2.21) and (2.22) into account, for any distinct set $(i, j, k, s)$ we have three independent permutations of indices. According to the symmetry (2.23) one of these three is dependent of the other two, and is redundant. The number of all such redundant sets is:

\[ \binom{n}{4}. \quad (2.26) \]

The number of independent components is then obtained by subtracting (2.25) and (2.26) from (2.24)

\[ N - \binom{m}{2} - \binom{n}{4} = \frac{1}{12} n^2 (n^2 - 1). \quad (2.27) \]

For a three dimensional manifold this means six independent components. We could visualize these symmetries by taking the indices as vertices of a tetrahedron, drawing symmetry axes from one fixed vertex to the barycenter of three other vertices, from the center of an edge to the center of the opposite edge and two symmetry planes as in Fig. 2.1.
Figure 2.1: Two axes and two planes of symmetry for indices of the Riemann tensor. A rotation around the axis that exchanges the positions of the indexed balls or a reflection w.r.t. the plane does not change the value of the tensor component. There are two other pairs of axes and two other pairs of planes of symmetry, but they are not shown here to avoid clutter.
The symmetry properties can be used to simplify computations in lower dimensions. As an example we compute all possible Riemann tensors in dimension two.

\[ R_{1111} = g_{11} R_{111}^1 + g_{21} R_{111}^2 \]
\[ = g_{11} (\Gamma_{111}^1 + \Gamma_{112}^1 + \frac{\partial}{\partial x^1} \Gamma_{111}^1 - \Gamma_{111}^1 \Gamma_{111}^1 - \Gamma_{112}^1 \Gamma_{112}^1 - \frac{\partial}{\partial x^1} \Gamma_{111}^1) \]
\[ + g_{21} (\Gamma_{111}^2 + \Gamma_{112}^2 + \frac{\partial}{\partial x^1} \Gamma_{111}^2 - \Gamma_{111}^2 \Gamma_{111}^2 - \Gamma_{112}^2 \Gamma_{112}^2 - \frac{\partial}{\partial x^1} \Gamma_{111}^2) = 0 . \]

(2.28)

This implies that also \( R_{2222} = 0 \). The next term is

\[ R_{1112} = g_{12} R_{111}^1 + g_{22} R_{111}^2 \]
\[ = g_{12} \cdot 0 + g_{22} \cdot 0 \]
\[ = 0 . \]

(2.29)

From this follows again by symmetry w.r.t. indices,

\[ R_{1121} = R_{1211} = R_{2111} = R_{2221} = R_{2212} = R_{2122} = R_{1222} = 0 . \]

(2.30)

Also

\[ R_{1122} = R_{2211} = 0 . \]

(2.31)

The only non-zero terms are

\[ R_{1212} = R_{2121} = -R_{1221} = -R_{2112} . \]

(2.32)

**Definition 2.9.** The Ricci tensor \( R_{ik} \) is defined as

\[ R_{ik} = g^{il} R_{ijkl} . \]

(2.33)

Again, in dimension two, the coefficients of Ricci tensors are simply

\[ R_{11} = g^{22} R_{1212} \]
\[ R_{12} = -g^{21} R_{1212} \]
\[ R_{21} = -g^{12} R_{1212} \]
\[ R_{22} = g^{11} R_{1212} . \]

(2.34)
So, the Ricci tensor reduces to
\[
R_{ij} = \frac{R_{1212}}{\det(g)} g^{ij}.
\] (2.35)

where \(g^{ij}\) is defined by \(g^{ik}g_{kj} = \delta^i_j\), i.e. is the inverse of \(g_{ij}\) w.r.t. orthogonal basis (which in this thesis is taken to be the identity matrix).

**Definition 2.10.** Let \(\sigma\) be a two dimensional subspace of the tangent space \(T_pM\). Let \(X, Y \in \sigma\) be two linearly independent vectors. We define the sectional curvature of \(\sigma\) at \(p\) as
\[
K(X, Y) = \frac{\langle R(X,Y)X,Y \rangle}{\langle |X|^2|Y|^2 - \langle X,Y \rangle^2 \rangle}.
\] (2.36)

Sectional curvature is thus defined by Riemann curvature and Riemann curvature on the other hand can be expressed as a sum of sectional curvatures ([53] p.137).

The sectional curvature does not require the input-vectors to be orthogonal, as long as they span a plane i.e. are not parallel. This can be seen by simple computation
\[
X \perp_{g_{ij}} Y, \quad W = \alpha X + \beta Y, \quad \alpha \neq 0, \quad \beta \neq 0,
\] (2.37)
and working out as follows
\[
\frac{\langle R(X,W)X,W \rangle}{\langle |X|^2|W|^2 - \langle X,W \rangle^2 \rangle} = \frac{\langle R(X,\alpha X + \beta Y)X,\alpha X + \beta Y \rangle}{\langle |X|^2|\alpha X + \beta Y|^2 - \langle X,\alpha X + \beta Y \rangle^2 \rangle} = \frac{\langle R(X,Y)X,Y \rangle}{\langle |X|^2|Y|^2 - \langle X,Y \rangle^2 \rangle}.
\] (2.38)

For parametrized surfaces \((z = f(x,y))\) embedded in \(\mathbb{R}^3\), the Gaussian curvature \(G\) is defined as:
\[
G = \frac{\det(II)}{\det(I)} = \frac{f_{xx}f_{yy} - (f_{xy})^2}{(1 + f_x^2 + f_y^2)^2},
\] (2.39)
where \(I\) and \(II\) are the so-called first and second fundamental forms. The first fundamental form is equal to the metric tensor and the second fundamental form is equal to the Hessian divided by the square root of the determinant of the metric tensor. In dimension two
\[
G = R_{1212} = K\left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}\right).
\] (2.40)
In dimension \( n = 3 \), there are infinitely many planes containing a given tangent vector. To obtain a scalar value associated with a given vector involving sectional curvature, the mean of all sectional curvatures can be taken. This is the Ricci curvature (or mean curvature).

**Definition 2.11.** The Ricci scalar \( R \) is defined as follows

\[
R = R_{ik}g^{ik}.
\]  

(2.41)

This is related to the scalar curvature \( K \) (in dimension \( n \)) as follows

\[
K = \frac{1}{n(n-1)} R.
\]  

(2.42)

It is possible to express the Riemann tensor in dimensions two and three in terms of the Ricci scalar \( R \), the Ricci tensor \( R_{ij} \) and the metric tensor \( g_{ij} \). In dimension two there is only one independent Riemann tensor component and for any non-zero vector there is a unique independent vector \( V^\perp \) up to a scalar factor, which relates to \( V \) as

\[
g_{ij}V^i(V^\perp)^j = 0,
\]  

(2.43) and \( R_{ijkl} \) can be expressed in terms of \( R \) and \( g \). Using equation (2.33), we have

\[
R = g^{11}g^{22}R_{1212} + g^{12}g^{21}R_{1221} + g^{21}g^{12}R_{2112} + g^{22}g^{11}R_{2121}.
\]  

(2.44)

Since

\[
R_{1221} = R_{2112} = -R_{1212}, \quad \text{and} \quad R_{2121} = R_{1212},
\]  

(2.45) this becomes

\[
R_{1212} = \frac{1}{2} \det(g)R.
\]  

(2.46)

In a general form \( R_{ijkl} \) can be written as

\[
R_{ijkl} = \frac{1}{2} R(g_{ik}g_{jl} - g_{ij}g_{kl}).
\]  

(2.47)

In dimension three, we have six independent Riemann tensor components. Each of these can also be represented in terms of the metric- and the Ricci
2.3. Geodesics With Fixed Initial Point and Direction

The Ricci tensor and Riemann tensor are equivalent in the sense that knowing the independent Riemann components, we can compute the Ricci tensor and vice versa. Without proof, we mention the formula with which one can express Riemann tensor components in terms of Ricci tensor components and Ricci scalar in case $n = 3$. We refer to [61] for details.

\[
R_{ijkl} = (g_{ik}R_{jl} - g_{il}R_{jk} - g_{jk}R_{il} + g_{jl}R_{ik}) - \frac{R}{2}(g_{ik}g_{jl} - g_{il}g_{jk}). \quad (2.48)
\]

There are many books on Riemannian geometry. Some of them are more theoretical [19],[59],[53],[64] and some more oriented in applications to physics [67],[96].

2.3 Geodesics With Fixed Initial Point and Direction

Let $\gamma(t) : I \rightarrow M$, be a smooth parameterized curve on the manifold $M$ and let a vector field $V$ along the curve $\gamma$ be extendible (Definition 8.4). The tangent vector of the curve is $\dot{\gamma}(t) = \dot{\gamma}^i(t)\partial_i$ in the induced coordinates on the tangent space $T_{\gamma(t)}M$. The covariant derivative of the vector field $V$ along $\gamma$ is then

\[
\nabla_{\dot{\gamma}^i\partial_i} V = V^j \nabla_{\dot{\gamma}^i\partial_i} \partial_j + (\dot{\gamma}^i\partial_i V^j) \partial_j = (\dot{\gamma}^i V^j \Gamma^k_{ij} + \dot{V}^k) \partial_k \overset{\text{def.}}{=} D_t V. \quad (2.49)
\]

Substituting $\dot{\gamma}$ in place of $V$, in the equation (2.49) gives

\[
\nabla_{\dot{\gamma}^i\partial_i} (\dot{\gamma}^j \partial_j) = (\dot{\gamma}^i \dot{\gamma}^j \Gamma^k_{ij} + \dot{\gamma}^k) \partial_k \overset{\text{def.}}{=} D_t \dot{\gamma}^j \partial_j. \quad (2.50)
\]

In analogy to the straight lines in a Euclidean space, a geodesic is defined to be a curve whose acceleration is zero. Thus a geodesic $\gamma$ satisfies the following equation

\[
\ddot{\gamma}^k + \Gamma^k_{ij} \dot{\gamma}^i \dot{\gamma}^j = 0. \quad (2.51)
\]
Definition 2.12. Let $M$ be a differentiable manifold with a linear connection $\nabla$. A vector field $V$ along a curve $\gamma : I \rightarrow M$ is called parallel, when $\nabla_\gamma V = 0$.

Geodesics are thus also called autoparallel curves. An important property of a geodesic is the following. If $\gamma : I \rightarrow M$ is a geodesic then

$$\frac{d}{dt} \langle \dot{\gamma}, \dot{\gamma} \rangle = 2\langle \nabla_{\dot{\gamma}} \dot{\gamma}, \dot{\gamma} \rangle = 0 . \quad (2.52)$$

This means that the length of the tangent vector $\dot{\gamma}$ is constant. If we look at the arclength of $\gamma$ between points $\gamma(t_0)$ and $\gamma(t)$

$$s(t) = \int_{t_0}^{t} |\dot{\gamma}| \, dt = c(t - t_0) , \quad (2.53)$$

we see that the parameter of a geodesic is proportional to arc length.

### 2.4 Geodesic Deviation

In this section, we assume again that every manifold $M^n$ is a Riemannian manifold (Definition 8.1), with metric tensor $g_{ij}$ and associated affine connection. From now on, we also assume that a manifold $M^n$ is compact. This is practical, since in a compact space every Cauchy-sequence converges and the metric is complete. The Hopf-Rinow theorem guarantees that if a Riemannian manifold is a complete metric space, then it is also geodesically complete i.e. any two points on $M$ can be connected by a minimal geodesic. In the first subsection we define the so-called exponential mapping in detail, referring to [19], since it forms a basis for the applications later.

#### 2.4.1 Exponential Mapping

Theorem 2.13 (Existence and uniqueness of geodesics). Let $M$ be a Riemannian manifold with a linear connection. Given any $p \in M$, any $V \in T_pM$ and any $t_0 \in \mathbb{R}$, there exists an open interval $I \subset \mathbb{R}$ s.t. $t_0 \in I$ and a geodesic $\gamma : I \rightarrow M$ such that $\gamma(t_0) = p$, $\dot{\gamma}(t_0) = V$. Any pair of these geodesics agree on their common domain.
For proof, see [58], p.58 and the references therein.

In this section, we denote a geodesic \( \gamma \) starting from point \( p \), with initial tangent vector \( V \), at \( t \) as
\[
\gamma(t, p, V).
\] (2.54)

In the definition of the exponential map, we use the following result that is referred to as homogeneity of a geodesic

**Lemma 2.14.** If the geodesic \( \gamma(t, p, V) \) on \( M \) is defined on the interval \((-\varepsilon, \varepsilon)\), then the geodesic \( \gamma(ct, p, cV) \), \( (c \in \mathbb{R}, \ c > 0) \) is defined on the interval \((-\frac{\varepsilon}{c}, \frac{\varepsilon}{c})\). And
\[
\gamma(t, p, cV) = \gamma(ct, p, V) .
\] (2.55)

**Proof.**

Let \( h : (-\frac{\varepsilon}{c}, \frac{\varepsilon}{c}) \) be a curve defined by
\[
h(t) = \gamma(ct, p, V) .
\] (2.56)

For such a curve we have \( h(0) = p \) and \( \dot{h}(0) = cV \). Also
\[
\dot{h}(t) = c\dot{\gamma}(ct, p, V) ,
\] (2.57)

and assuming an extendible vector field
\[
\nabla_{c\dot{\gamma}(ct, p, V)} (c\dot{\gamma}(ct, p, V)) = c^2 \nabla_{\dot{\gamma}(ct, p, V)} (\dot{\gamma}(ct, p, V)) = 0 .
\] (2.58)

Then \( h(t) \) is a geodesic through \( p \) with initial velocity \( cV \) and by uniqueness
\[
\gamma(ct, p, V) = \gamma(t, p, cV) .
\] (2.59)

From the previous lemma it follows that there exists an \( \varepsilon_0 \), such that for any vector \( V \in T_pM \) with \( ||V|| = 1 \), \( \gamma_V \) is defined in \( t \in [0, \varepsilon_0] \). And from this in turn we have that for any \( W \in T_pM \), \( ||W|| = \varepsilon_0 \), \( \gamma_W \) is defined at least on an interval \([0, 1]\).

We define a subset \( \mathcal{E} \) in \( TM \).
\[
\mathcal{E} = \{ V \in TM \mid \gamma_V \text{ is defined on an interval } I \text{ containing } [0, 1] \} .
\] (2.60)

Then the exponential map is defined as
\[
\exp : \mathcal{E} \rightarrow M, \quad \exp(V) = \gamma_V(1) .
\] (2.61)
The restriction of the exponential map to set $E_p = \mathcal{E} \cap T_pM$ is denoted as $\exp_p$. The point $\exp_p(V)$ is thus the point on a manifold that we reach by proceeding on the geodesic, initially in the direction $V$ for time 1 and with velocity $|V|$. Since there is the trade-off by Lemma 2.14, this is equivalent to saying that the point $\exp_p(V)$ is the point on a manifold that we reach by proceeding on the geodesic, initially in the direction $V$ for time $|V|$ and with velocity $\frac{V}{|V|}$.

According to the theorem (2.13), there is an interval that $\exp_p(V)$ exists. Due to same theorem, we know that there also exists an interval $(-\varepsilon, \varepsilon)$, such that $\exp_p(U)$ is defined for

$$U = tV(s), \quad 0 \leq t \leq 1, \quad -\varepsilon < s < \varepsilon,$$  \hfill (2.62)

where $V(s)$ is a curve in $T_pM$, for which $V(0) = V$ and $V'(0) = W$ and $|V(s)|$ is constant. The variation of $s$ does not increase the length of the vector $tV(s)$, i.e. the parameter $s$ does not contribute into the direction of $tV$ and $\frac{d}{ds} tV(s)$ is for every $0 \leq t \leq 1$, orthogonal to the vector $tV$ itself. We define a parametrized surface $f$ from tangent space $T_pM$ to the manifold as follows

$$f(t, s) = \exp_p tV(s), \quad 0 \leq t \leq 1, \quad -\varepsilon < s < \varepsilon.$$  \hfill (2.63)

This defines a surface made of a continuum of geodesics (see Fig. 2.2).

Let $V \in T_pM$ and $W \in T_V (T_pM) \approx T_pM$. We consider a subset of the previous surface, namely we fix $t = 1$. Now $f(1, s)$ is a curve segment on the manifold, going through the point $\exp_p(V)$, whose tangent vector at this point we want to compute. The derivative of the exponential mapping $\exp_p(tV)$ in the direction of $W$ at $t = 1$, i.e. the tangent vector of the curve $f(1, s)$ at $s = 0$ is:

$$(d \exp_p)_V W = \frac{\partial f}{\partial s}(1, 0).$$  \hfill (2.64)

The absolute value $|(d \exp_p)_V(W)|$ measures the length of this vector. If we express the derivative (2.64) as a function of $t$ in a small neighborhood of $p$, we have

$$(d \exp_p)_tV tW = \frac{\partial f}{\partial s}(t, 0).$$  \hfill (2.65)

As is shown shortly, the vector field (2.65) is an example of a so called Jacobi field. Using properties of this field one can compute an estimate of
the length of (2.65). Recall here the definition $D_t V := \nabla_\gamma V$ for a vector $V$ at point $\gamma(t)$ (2.50).

**Definition 2.15.** The equation

$$D_t D_t J(t) + R(\dot{\gamma}, J(t))\dot{\gamma} = 0,$$  \hspace{1cm} (2.66)

is called the Jacobi equation and a vector field $J(t)$ along a geodesic $\gamma$, satisfying the Jacobi equation is called a Jacobi field.

We use the fact that

$$D_t \frac{\partial f}{\partial t} = 0,$$ \hspace{1cm} (2.67)

since with fixed $s_0$, $f(t,s_0)$ is a geodesic, and a lemma (for details see [19] p. 98.), which says that for a parameterized surface $f(s,t)$ and a vector field $V(s,t)$ the following holds

$$D_t D_s V - D_s D_t V = R\left(\frac{\partial f}{\partial s}, \frac{\partial f}{\partial t}\right)V.$$ \hspace{1cm} (2.68)

From (2.67) and (2.68)

$$D_s \left( D_t \frac{\partial f}{\partial t} \right) = 0$$ \hspace{1cm} (2.69)

and since $R(X,Y)Z = -R(Y,X)Z$ and $D_t \frac{\partial f}{\partial s} = D_s \frac{\partial f}{\partial t}$ in local coordinates for orthogonal vectors $\frac{\partial}{\partial t}$ and $\frac{\partial}{\partial s}$ ([19] p. 68)

$$D_s(D_t \frac{\partial f}{\partial t}) = D_t D_s \frac{\partial f}{\partial t} - R\left(\frac{\partial f}{\partial s}, \frac{\partial f}{\partial t}\right)\frac{\partial f}{\partial t}$$

$$= D_t D_t \frac{\partial f}{\partial s} + R\left(\frac{\partial f}{\partial t}, \frac{\partial f}{\partial s}\right)\frac{\partial f}{\partial t}$$

$$= 0.$$ \hspace{1cm} (2.70)

Abbreviating $\frac{\partial f}{\partial s}(t,0)$ as $J(t)$ we deduce from the equation (2.70) that

$$D_t D_t J(t) + R(\dot{\gamma}, J(t))\dot{\gamma} = 0,$$ \hspace{1cm} (2.71)

and $J(t) := \frac{\partial f}{\partial s}(t,0)$ indeed satisfies the Jacobi equation.

We consider a Jacobi field

$$J(t) = (d \text{exp}_p)_t V(tW),$$ \hspace{1cm} (2.72)
for which \( |V| = 1, |W| = 1 \), and \( \langle V, W \rangle = 0 \), and calculate the Taylor expansion of \( |J(t)|^2 \) around \( t = 0 \).

We know the initial conditions \( J(0) = 0 \), and \( J'(0) = W \). Also we know from equation (2.71) and the bilinearity of Riemann curvature, that

\[
J''(0) = -R(\gamma'(0), J(0))\gamma'(0) = 0.
\]

Using these facts, we can calculate the inner products:

\[
\begin{align*}
\langle J, J \rangle(0) & = 0 \\
\langle J, J' \rangle(0) & = 0 \\
\langle J, J'' \rangle(0) & = 2\langle J'', J \rangle(0) + 2\langle J', J' \rangle(0) = 2\langle W, W \rangle = 2. \tag{2.73} \\
\langle J, J''' \rangle(0) & = -8\langle J', R(\gamma', J')\gamma' \rangle(0) = -8\langle R(V, W)V, W \rangle(0).
\end{align*}
\]

For details concerning the fourth derivative see [19] p.115. From these, we can compute the fourth order Taylor expansion about \( t = 0 \):

\[
|J(t)|^2 = t^2 - \frac{1}{3} \langle R(V, W)V, W \rangle t^4 + O(t^5). \tag{2.74}
\]

From which we calculate:

\[
|J(t)| = t - \frac{1}{6} \langle R(V, W)V, W \rangle t^3 + O(t^4). \tag{2.75}
\]

As we can see in Fig. 2.2(b), the geodesics \( l_1, l_2 \) in \( T_pM \) deviate from each other at rate \( t \). Their images (the geodesics) under exponential mapping on the manifold deviate from each other approximately at rate \( |J(t)| \).

Here \( (|V|^2|W|^2 - \langle V, W \rangle^2) = 1 \) and the \( \langle R(V, W)V, W \rangle \) is equal to the sectional curvature (2.36) with respect to plane \( \sigma \) generated by vectors \( V, W \). From

\[
|J(t)|^2 = t^2 - \frac{1}{3} K(p, \sigma) t^4 + R(t), \tag{2.76}
\]

we get

\[
|J(t)| = t - \frac{1}{6} K(p, \sigma) t^3 + R_1(t), \tag{2.77}
\]

where \( \lim_{t \to 0} \frac{R_1(t)}{t^3} = 0 \). Compared to the angle between their tangent vectors in \( T_pM \), two geodesics deviate more if \( K(p, \sigma) < 0 \) and less if \( K(p, \sigma) > 0 \).
2.4.2 Geodesic Deviation

We have a good approximation (2.77) to the local deviation of geodesics with respect to the plane generated by the vectors \( V, W \). This result requires one to specify a unit vector \( W \perp V \), where \( V \) is the tangent of the geodesic. The interesting geometric quantity in (2.77) is the inner product \( \langle R(V, W) V, W \rangle \). To obtain an unambiguous scalar measure of geodesic deviation w.r.t. \( V \), we take an average of \( \langle R(V, W_i) V, W_i \rangle \) over every orthogonal plane generated by \( V \) and \( W_i \). In dimension \( n = 2 \) there is only one independent unit vector \( W \) orthogonal to a vector \( V \) and only two linear combinations of \( W \) with length 1. So the average of the geodesic deviation is:

\[
\text{average deviation} = \frac{1}{2} \left( \langle R(V, W) V, W \rangle + \langle R(V, -W) V, -W \rangle \right) = \langle R(V, W) V, W \rangle .
\]  

(2.78)

When \( n = 3 \), we obtain the average as follows. Pick any two independent orthonormal vectors \( W_1, W_2 \perp V \). Then take the average over all linear
combinations of \( W_1, W_2 \), that are of unit length:

\[
\text{average deviation} = \frac{1}{2\pi} \int_0^{2\pi} \langle R(V, \cos \theta W_1 + \sin \theta W_2)V, \cos \theta W_1 + \sin \theta W_2 \rangle d\theta \\
= \frac{\pi}{2\pi} \left( \langle R(V, W_1)V, W_1 \rangle + \langle R(V, W_2)V, W_2 \rangle \right) \\
= \frac{1}{2} \sum_{i=1}^{2} \langle R(V, W_i)V, W_i \rangle.
\]

(2.79)

The cases \( n > 3 \) can be calculated similarly. Below we show that the result does not depend on the choice of orthonormal basis \( \{W_i\}_{i=1}^{n-1} \) of \( V^\perp \).

The obtained quantities are actually the Ricci curvatures \( Ric_p \) in dimension 2 and 3 respectively, since Ricci curvature is defined as

\[
Ric_p(V) = \frac{1}{n-1} \sum_{i=1}^{n-1} \langle R(V, W_i)V, W_i \rangle.
\]

(2.80)

In terms of a local coordinate basis, we may write the Ricci curvature (2.80):

\[
Ric_p(V) = \frac{1}{n-1} \sum_{h=1}^{n-1} \langle V^i W_h^j V^k R_{ijk}^l \frac{\partial}{\partial x^l}, W_h^m \frac{\partial}{\partial x^m} \rangle \\
= \frac{1}{n-1} \sum_{h=1}^{n-1} V^i V^k W_h^m W_h^j R_{ijk}^l g_{lm} \\
= \frac{1}{n-1} \sum_{h=1}^{n-1} V^i V^k W_h^m W_h^j R_{ijkm}
\]

(2.81)

where \( V = V^i \frac{\partial}{\partial x_i} \), and so on.

**Lemma 2.16.** Let \( \{W_1, W_2, \ldots, W_n\} \) constitute an orthonormal basis in the \( n \)-dimensional tangent space \( T_p M \), then

\[
\sum_{h=1}^{n} W_h^j W_h^m = g^{jm}.
\]

(2.82)
Proof.
Since \( \{W_1, W_2, \ldots, W_{n-1}, W_n\} \) is an orthonormal basis in the tangent space,
\[
g_{ij} W^i_k W^j_l = \delta_{kl} .
\]
We denote by \( W \) the matrix
\[
W = \begin{pmatrix} W_1 & W_2 & \cdots & W_{n-1} & W_n \end{pmatrix}
\]
having the basis vectors as columns. Then
\[
\sum_{h=1}^{n} W^j_h W^m_h = (W \cdot W^T )_{jm} .
\]
Denoting the metric tensor as \( G \), from (2.83) we have
\[
W^T GW = I .
\]
By simple manipulation of (2.86) we get
\[
G = (WW^T )^{-1} ,
\]
from which we obtain
\[
G^{-1} = WW^T ,
\]
and the equivalence of individual components in (2.82) follows. Since lemma (2.16) holds for any orthonormal basis it holds also for a basis
\[
W = \begin{pmatrix} W_1 & W_2 & \cdots & W_{n-1} & V \end{pmatrix}
\]
with fixed last component. Then we have
\[
\sum_{h=1}^{n-1} W^j_h W^m_h = g^{jm} - V^j V^m .
\]
Substitution of this in (2.81) yields
\[
Ric_p(V) = \frac{1}{n-1} V^i V^k (g^{jm} - V^j V^m) R_{ijkm}
\]
\[
= \frac{1}{n-1} V^i V^k R_{ijkm} g^{mj} - \frac{1}{n-1} V^i V^k V^j V^m R_{ijkm}
\]
\[
= \frac{1}{n-1} V^i V^k R_{ijkm} g^{mj}
\]
\[
= \frac{1}{n-1} V^i V^k R_{ik} ,
\]
by definition (2.9) and because the latter summand in the second row vanishes by virtue of the symmetries of Riemann tensor. In chapter 3, we interpret the Ricci curvature as the relative acceleration of the flux of geodesics in a given direction.
2.5 Geodesics With Fixed End Points

On a Riemannian manifold, a minimal geodesic between two fixed points is a curve whose length with respect to the arc length parameter is the unique minimum of all possible lengths. This means that if \( x(t) \) is a shortest path, then the following functional is minimized:

\[
L(x) = \int_{t_0}^{t_1} \sqrt{g_{ij}(x(t))\dot{x}^i(t)\dot{x}^j(t)} \, dt.
\]  \tag{2.92}

It can be shown that minimal geodesics simultaneously minimize energy

\[
E(x) = \int_{t_0}^{t_1} \frac{1}{2} g_{ij}(x(t))\dot{x}^i(t)\dot{x}^j(t) \, dt.
\]  \tag{2.93}

For the rest of the chapter we suppress the parameters of the metric tensor from notation and put

\[
g_{ij} := g_{ij}(x(t)).
\]  \tag{2.94}

To derive the equations that describe the extremals of these functionals, we consider the following linear functional, the so called fundamental integral

\[
\Phi[x] = \int_{t_0}^{t_1} L(t, x^i(t), \dot{x}^i(t)) \, dt,
\]  \tag{2.95}

where the function \( L(t, x^i(t), \dot{x}^i(t)) \) is called the Lagrangian. The Lagrangian is a scalar valued function of \( 2n + 1 \) variables. In classical mechanics it describes the difference between kinetic and potential energies.

**Definition 2.17.** For a linear differentiable functional \( \Phi \), we can split the increment

\[
\Delta \Phi[x] = \Phi[x + h] - \Phi[x],
\]

into two parts

\[
\Phi[x + h] - \Phi[x] = L[h] + \varepsilon \|h\|, \\
\]

where \( L[h] \) is the principal linear part called the (first order) variation and \( \lim_{h \to 0} \frac{\varepsilon \|h\|}{\|h\|} = 0 \).
Theorem 2.18. A necessary condition for the functional $\Phi[x]$ to have an extremum for $x(t) = \gamma(t)$ is that its variation is zero.

The vanishing of the variation of $\Phi[x]$ is in turn equivalent to the Euler-Lagrange equations [39][64]:

$$\frac{d}{dt} \left( \frac{\partial L}{\partial \dot{x}^j} \right) - \frac{\partial L}{\partial x^j} = 0 . \quad (2.96)$$

For the energy functional (2.93) we have

$$\frac{d}{dt} \left( \frac{\partial L}{\partial \dot{x}^i} \right) - \frac{\partial L}{\partial x^i} = 2g_{ln} \ddot{x}^n + \frac{\partial g_{lj}}{\partial x^n} \dot{x}^i \dot{x}^j + \frac{\partial g_{m}^{l}}{\partial x^m} \ddot{x}^m \dot{x}^k - \frac{\partial g_{ij}}{\partial x^l} \dot{x}^i \dot{x}^j$$

$$= 2g_{ln} \ddot{x}^n + \left( \frac{\partial g_{lj}}{\partial x^n} + \frac{\partial g_{il}}{\partial x^j} - \frac{\partial g_{ij}}{\partial x^l} \right) \dot{x}^i \dot{x}^j$$

$$= 0 . \quad (2.97)$$

From the last equality, we get

$$\ddot{x}^n + \Gamma^x_{ij} \dot{x}^i \dot{x}^j = 0, \quad (2.98)$$

the Riemannian geodesic equation.

Lemma 2.19. Let us assume that the extremum $x(t)$ has constant velocity. Then from the Euler-Lagrange equations for expression (2.92) follow that this extremum also satisfies the Riemannian ODE for geodesics.

Proof.

The Euler-Lagrange equations with Lagrangian $L = \sqrt{g_{ij} \ddot{x}^i \ddot{x}^j}$, using the abbreviation $G = g_{ij} \dot{x}^i \dot{x}^j$ is

$$E_k(L) = \frac{d}{dt} \left( \frac{\partial L}{\partial \dot{x}^k} \right) - \frac{\partial L}{\partial x^k}$$

$$= -G^{-\frac{3}{2}} \left( \frac{\partial g_{lj}}{\partial x^n} \dot{x}^l \dot{x}^j + 2g_{ni} \ddot{x}^l \dot{x}^n \right) g_{sk} \dot{x}^k + G^{-\frac{1}{2}} \frac{\partial g_{r}}{\partial x^p} \dot{x}^p \dot{x}^r$$

$$+ G^{-\frac{1}{2}} g_{sk} \ddot{x}^s - \frac{1}{2} G^{-\frac{1}{2}} \frac{\partial g_{tm}}{\partial x^k} \dot{x}^t \dot{x}^m$$

$$= 0 . \quad (2.99)$$

This can be simplified by multiplying both sides with $G^{3/2}$ and $g^{lk}$. 
\[
E_k(L) = \dddot{x}^l + \Gamma_{ij}^l \dot{x}^i \dot{x}^j - \frac{1}{G} \dot{x}^l (2g_{rh} \dot{x}^r \dot{x}^h + \frac{\partial g_{nq}}{\partial x^s} \dot{x}^n \dot{x}^q \dot{x}^s) \\
= \dddot{x}^l + \Gamma_{ij}^l \dot{x}^i \dot{x}^j
\]
\[\text{(2.100)}\]

since \( \frac{\partial L}{\partial t} \) vanishes and we obtain the usual equation for a geodesic.

To find the shortest paths between a fixed point \( x_0 = x(0) \) and an arbitrary point \( p \) on a compact manifold, we may consider the following heuristic approach.

1. Calculate the distance function \( S(x, t) : \mathbb{R}^{n+1} \to \mathbb{R} \):

\[
x(t)=p \\
S(p, t) = \min_{x(t)} \int_0^t L(t, x, \dot{x}) dt,
\]

that assigns to every point \( p \) the minimal length (2.92) or energy (2.93).

2. Follow a geodesic (a characteristic of the distance function) from \( p \) to \( x_0 \) by proceeding along the shortest (Euclidean) paths from a level set to another.

In a discrete grid the first part is typically solved numerically using level set methods [76] [85]. The latter part corresponds to solving the canonical equations which we introduce shortly.

It can be shown [38] that the minimal distance function \( S \) satisfies the so-called Hamilton-Jacobi equation

\[
\frac{\partial S}{\partial t} + H \left( t, x^i, \frac{\partial S}{\partial x^i} \right) = 0.
\]

\[\text{(2.102)}\]
When solving the minimal distance function, there are two different cases, depending on whether the Lagrangian is homogeneous or not. In the following we briefly discuss these two cases.

### 2.5.1 The Hamilton-Jacobi Equation for a Non-Homogeneous Integrand

First we define the momentum $p$ of a system from the Lagrangian function

$$p_j = \frac{\partial L(t, x^h, \dot{x}^h)}{\partial \dot{x}^j}. \quad (2.103)$$

The Hamiltonian function $H$ is obtained from the Lagrangian $L$ as follows

$$H(t, x^h, p_h) = -L \left( t, x^h, \phi^h(t, x^l, p_l) \right) + p_j \phi^j(t, x^h, p_h), \quad (2.104)$$

where $\phi^h(t, x^l, p_l)$ is just $\dot{x}^h$ expressed as function of $t, x^l$, and $p_l$. According to the inverse function theorem this can be done, when the mapping

$$\dot{x}^j \rightarrow \frac{\partial L}{\partial \dot{x}^j} \quad (2.105)$$

is a bijection for every $j$, that is

$$\det \left( \frac{\partial^2 L}{\partial \dot{x}^i \partial \dot{x}^j} \right) \neq 0. \quad (2.106)$$

This transformation from $L$ to $H$ is called the Legendre transformation, and whereas $L$ is a function of spatial coordinates and velocity, $H$ is a function of spatial coordinates and momentum. The variables $x_i, p_i$ are the so called canonical variables. This name for the variables comes from the fact that they allow the reduction of $n$ second order Euler-Lagrange equations to $2n$ concise first order equations involving the Hamiltonian. We do this reduction in (2.111).

In case of our special Lagrangian (2.93), (2.103) becomes

$$p_j = g_{ij} \dot{x}^i, \quad (2.107)$$

and thus

$$\dot{x}^j = g^{ij} p_i, \quad (2.108)$$
where \( g^{ij} \) is the inverse of \( g_{ij} \).

The Hamiltonian function (2.104) becomes

\[
H = -\frac{1}{2} g_{kl} g_{mk} p_m p_n + p_j g^{ij} p_i \\
= \frac{1}{2} g^{ln} p_l p_n. 
\]  

(2.109)

Using (2.104), the minimization problem in (2.93) can be transformed into an equivalent problem of minimizing

\[
\int_{t_0}^{t_1} -H(t, x^k, p^k) + p_j \dot{x}^j(t, x^k, p^k) dt, 
\]

(2.110)

with the respective Euler-Lagrange equations:

\[
\frac{\partial H}{\partial t} = -\frac{\partial L}{\partial t}, \quad \frac{\partial H}{\partial x^l} = -\frac{\partial L}{\partial x^l}, \quad \frac{\partial H}{\partial p^l} = \dot{x}^l. 
\]

(2.111)

Using the last two equations in (2.111), we can transform the Euler equations for Lagrangian \( L \)

\[
\dot{x}^i = \frac{dx}{dt}, \quad \frac{\partial L}{\partial x^i} = \frac{d}{dt} \left( \frac{\partial L}{\partial \dot{x}^i} \right), 
\]

(2.112)

to the canonical equations for Hamiltonian \( H \)

\[
\dot{x}^i = \frac{\partial H}{\partial p_i}, \\
\dot{p}_i = -\frac{\partial H}{\partial x^i}. 
\]

(2.113)

With a Lagrangian as (2.93) these are of course

\[
\dot{x}^i = g^{ij} p_j, 
\]

(2.114)

and

\[
\dot{p}_i = -\frac{1}{2} \frac{\partial g^{ij}}{\partial x^k} p_i p_j. 
\]

(2.115)
2.5. Geodesics With Fixed End Points

The Hamilton-Jacobi equation (2.102) is then

\[
\frac{\partial S}{\partial t} + \frac{1}{2} g^{ij} \frac{\partial S}{\partial x^i} \frac{\partial S}{\partial x^j} = 0.
\]  

(2.116)

The initial condition for this equation being

\[
S(x, 0) = \begin{cases} 
0 & \text{if } x = x_0, \\
\infty & \text{otherwise}
\end{cases}
\]  

(2.117)

2.5.2 The Hamilton-Jacobi Equation for a Homogeneous Integrand

The homogeneity of the integrand \( L(t, x, \dot{x}) \) means

\[
L(t, x, \lambda \dot{x}) = \lambda L(t, x, \dot{x}).
\]  

(2.118)

For the fundamental integral to be parameter invariant, it means that under a monotonous change of parameter \( \tau = f(t) \)

\[
\int_{t_1}^{t_2} L(t, x, \dot{x}) dt = \int_{f(t_1)}^{f(t_2)} L(\tau, x, \dot{x}) d\tau.
\]  

(2.119)

For a homogeneous integrand, this is indeed the case since

\[
\int_{t_1}^{t_2} L(t, x, \dot{x}) dt = \int_{f(t_1)}^{f(t_2)} L(\tau, x, \dot{x}) f'(t) d\tau = \int_{f(t_1)}^{f(t_2)} L(\tau, x, \dot{x}) d\tau.
\]  

(2.120)

The homogeneity implies

\[
\frac{d}{d\lambda} (L(t, x, \lambda \dot{x})) = \left( \frac{d}{d\lambda} \lambda L(t, x, \dot{x}) \right),
\]  

(2.121)

i.e.

\[
L_{\dot{x}i} x^i = L.
\]  

(2.122)

Further, taking once more derivatives w.r.t. \( \dot{x} \)

\[
L_{\dot{x}i} \dot{x}^i + L_{\dot{x}j} = L_{\dot{x}j},
\]  

(2.123)
resulting in

\[ L_{\dot{x}^i,\dot{x}^i} x^i = 0. \tag{2.124} \]

From this follows that the determinant of the Jacobian of mapping

\[ \dot{x}^j \rightarrow \frac{\partial L}{\partial \dot{x}^j}, \]

is

\[ \det \left( \frac{\partial^2 L}{\partial \dot{x}^i \partial \dot{x}^j} \right) = 0. \tag{2.125} \]

In (2.92), the Lagrangian is indeed homogeneous. This means that the mapping is not invertible and we have no means to express the variable \( \dot{x} \) as a function of \( p_j \). For example if we tried to calculate the Hamiltonian function \( H \) from a homogeneous Lagrangian

\[ (g_{ij}\dot{x}^i\dot{x}^j)^{1/2} \]

\[ H(t, x, p) = \dot{x}^i p_i - L(t, x, \dot{x}) = \dot{x}^i p_i - (g_{ij}\dot{x}^i\dot{x}^j)^{1/2} \]

\[ = g_{ij}\dot{x}^i\dot{x}^j (g_{kl}\dot{x}^k\dot{x}^l)^{-1/2} - (g_{ij}\dot{x}^i\dot{x}^j)^{1/2} \]

\[ = 0. \tag{2.126} \]

In [83], it is shown how the canonical variables for the homogeneous case can be found by means of the fundamental tensor in Finsler geometry. Thus by replacing the momentum \( p_i \) with

\[ \tilde{p}_i = \tilde{g}_{ij}(x, \dot{x})\dot{x}^j, \tag{2.127} \]

where

\[ \tilde{g}_{ij} = \frac{1}{2} \frac{\partial^2 L^2(x, \dot{x})}{\partial \dot{x}^i \partial \dot{x}^j}. \tag{2.128} \]

We return to this fundamental tensor in the context of Finsler geometry in chapter 4.

### 2.5.3 Asymptotically Homogeneous Case Study

We look at an asymptotically homogeneous integrand case

\[ \int_{t_0}^{t_1} \frac{1}{\alpha} \left( g_{ij}\dot{x}^i\dot{x}^j \right)^{\alpha/2} dt, \tag{2.129} \]
where $\alpha \to 1$. We observe that when $\alpha \neq 1$, the functional is not parameter invariant and the Hamilton-Jacobi function can be derived as in subsection 2.5.1.

We calculate $p_j$:

$$p_j = \frac{\partial L}{\partial \dot{x}^j} = \left(g_{ik}\dot{x}^i \dot{x}^k\right)^{\alpha/2-1} g_{ij} \dot{x}^j. \quad (2.130)$$

From this follows that

$$p_j \dot{x}^j = \left(g_{ij}\dot{x}^i \dot{x}^j\right)^{\alpha/2-1} g_{ij} \dot{x}^i \dot{x}^j = \left(g_{ij}\dot{x}^i \dot{x}^j\right)^{\alpha/2}. \quad (2.131)$$

It is not obvious how to solve $\dot{x}^j$ from this equation. Instead, we guess for a solution and see whether it satisfies the equation. Suppose

$$\dot{x}^j = g^{ij} p_i (g^{lk} p_l p_k)^\beta. \quad (2.132)$$

Substituting this solution candidate to equation (2.131), the left-hand side becomes

$$g^{ij} p_i (g^{lk} p_l p_k)^\beta = \left(g^{ij} p_i p_j\right)^{(\beta+1)}, \quad (2.133)$$

and the right-hand side is

$$\left(g_{ij} g^{kt} p_t g^{lj} p_l (g^{rq} p_r p_q)^{2\beta}\right)^{\alpha/2} = \left(g^{ij} p_i p_j\right)^{\alpha \beta+1} \quad (2.134)$$

Thus (2.132) is a solution to (2.131) provided

$$\alpha \beta + 1 = \beta + 1. \quad (2.135)$$

If $\alpha = 1$, then $\beta$ and therefore the solution (2.132) is not unique. If $\beta = 0$, then we must have $\alpha = 2$. 
"Isn’t it enough to locate cortical areas engaged in deception, wrath, introspection, empathy? Do we really have to worry about their connections? The answer is "yes", principally because the function of a complex system cannot be deduced from an inventory of its components.” M-Marsel Mesulam
3.1 Introduction

A human cerebral cortex consists of approximately 20 billion neurons [77]. Neurons can communicate with nearby neurons via their dendrites or axons, but dendrites rarely spread further away from the neuron than 0.1 $\mu$m [48]. With remote neurons located in different regions of brain they can only communicate via thin pathways called axons. Such nerve fibers connect for example the retina to the visual cortex and the motor cortex to the spine or the spine to lower limbs etc. In this thesis we consider only axons within the brain. One neuron has only one axon but an axon may branch to connect multiple neurons. The network of axons in the brain is what is called the brain white matter. At the scale of the measurement, the brain white matter is homogeneous in terms of its chemical composition and therefore the conventional magnetic resonance imaging (MRI) cannot distinguish the architecture of the fiber bundles therein [71]. Diffusion Tensor Imaging (DTI) is a non-invasive MRI-technique to study anatomical structures by inferring the average intra-voxel incoherent motion or Brownian motion of water molecules in tissue. This is based on the observation that in average, particles will diffuse more along elongated structures such as bundles of axons. See Fig. 3.1 for illustration.

A typical voxel size in DTI is $1 - 2$ mm, which is very much larger than a
3.1. Introduction

A typical diameter (less than 10 $\mu$m) of a single axon in white matter. The white matter however, has also coherent bundles of fibers with diameters equal to or greater than the voxel dimensions.

The MRI measurement is based on the fact that hydrogen protons, when placed in a magnetic field and excited by radio frequency (RF) pulses, absorb electromagnetic radiation and re-emit this energy subsequently during their relaxation back to the original equilibrium situation [72]. In DTI the strength of the magnetic field varies according to position, forcing the nuclei to precess with position dependent frequencies. Given some time $\Delta$ to freely diffuse, the Brownian motion of these molecules dephases the position-dependent MR signal and results in a loss of signal. Although typically only $0.003\%$ of all water molecules in a voxel contribute to the signal/Tesla, in a $2 \times 2 \times 2$ mm voxel, with 1 Tesla field, this is still around $8.0 \times 10^{17}$ water molecules [71] and sufficient for a significant signal.

Let $x(t)$ be the position of a molecule with diffusion coefficient $D$ at time $t$, undergoing Brownian motion. According to Einstein[26] and Smoluchowski[87], the average displacement in 1-D is

$$\sigma = |x(t) - x(0)| = \sqrt{2Dt} \quad (3.1)$$

DTI measures a sequence of one-dimensional average displacements along several fixed directions. For example the mean displacement of freely diffusing water molecules at temperature $37^\circ$ is

$$\sigma \approx 8\mu{m} \quad (3.2)$$
in 30 ms, which is a typical length of measurement.

In the original article of Stejskal and Tanner [88], the signal $S$ at time $t$ is modeled by a differential equation

$$\frac{dS(t)}{dt} = -D\gamma^2 \Delta f(t) S(0), \quad (3.3)$$

with a certain function $f$, that depends on the imaging sequence. With a typical sequence [88][71], this has a solution

$$S = S(0)e^{-\gamma^2 G^2 \delta^2 (\Delta - \frac{\delta}{2})D}, \quad (3.4)$$

which is the so-called Stejskal-Tanner equation, where $G$ the gradient strength ($\frac{T}{m}$=Tesla/meter), $\gamma$ is the gyromagnetic ratio ($\frac{2.675 \times 10^8}{sT}$) and $\delta$
the duration (s) of the gradient pulse. In practice we take the signal $S(0)$ to be a signal acquired without diffusion weighting i.e. with $G(t) = 0$.

Using a short hand notation

$$b = \gamma^2 G^2 (\delta)^2 (\Delta - \frac{\delta}{3})$$

(3.5)

we solve the so-called apparent diffusion coefficients (ADC) in several directions ($k = 1, \ldots, n$)

$$D_k = -\frac{1}{b} \ln \left( \frac{S_k}{S_0} \right).$$

(3.6)

DTI models the average diffusion profile with a symmetric quadratic surface that best (in $L_2$ w.r.t. the Frobenius norm) approximates the discrete set of directional measurements. This is the simplest smooth surface model for the diffusion profile that can also take some anisotropy into account. If the diffusion profile really is quadratic, we can express it as a function of vectors $b_k$, which encode the gradient strength $b = |b_k|$ and direction $\frac{b_k}{|b_k|}$

$$\ln \left( \frac{S_k}{S_0} \right) = -D_{ij} \frac{b^i_k b^j_k}{|b_k|}.$$  

(3.7)

Or, since we prefer unit vectors

$$y = (y^1, y^2, y^3) := (\sin \theta \cos \varphi, \sin \theta \sin \varphi, \cos \theta),$$

(3.8)

we get

$$\ln \left( \frac{S_k}{S_0} \right) = -|b_k| D_{ij} y^i_k y^j_k,$$

(3.9)

i.e.

$$D_{ij} y^i_k y^j_k = -\frac{1}{b} \ln \left( \frac{S_k}{S_0} \right).$$

(3.10)

At each voxel, the local diffusion profile is thus characterized by a diffusion tensor $D$

$$D = \begin{pmatrix} D_{11} & D_{12} & D_{13} \\ D_{21} & D_{22} & D_{23} \\ D_{31} & D_{32} & D_{33} \end{pmatrix}.$$ 

(3.11)

The diffusion tensor $D$ is real, symmetric and positive definite (SPD). This is because the diffusion is assumed to be symmetric w.r.t. the origin, and
3.1. Introduction

the diffusion coefficients to be real and positive. It is typically computed from a set of measurements by a least squares method as follows.

Let

\[ m = (m_1, \ldots, m_n)^T \]  

be the apparent diffusion coefficients (3.6) and

\[ y = (y_1, \ldots, y_n)^T \]  

be the \( n \) unit vectors i.e. the so-called gradient directions in which the measurements were taken. Denote the matrix components to be solved as

\[
\mathbf{d} = \begin{pmatrix}
D_{11} \\
D_{12} \\
D_{13} \\
D_{22} \\
D_{23} \\
D_{33}
\end{pmatrix}
\]  

and the corresponding monomials of gradient vector components of an \( i \)th measurement as

\[ M_i = \left((y_i^1)^2, y_i^1 y_i^2, y_i^1 y_i^3, (y_i^2)^2, y_i^2 y_i^3, (y_i^3)^2\right). \]  

Put

\[
M = \begin{pmatrix}
M_1 \\
M_2 \\
\vdots \\
M_n
\end{pmatrix},
\]  

and the tensor components can be solved by pseudo-inverse

\[ \mathbf{d} = \left(M^T M\right)^{-1} M^T m . \]  

For a SPD matrix, all eigenvalues \( \lambda_i \) are real and positive and all eigenvectors \( \mathbf{v}^i \) are orthogonal to each other. For illustrations of DTI tensors see Fig. 3.2. Before we proceed to the following applications of Riemannian geometry to DTI, we mention that some authors have already introduced Riemann geometry in the context of DTI \[75\] \[60\]. These first papers, propose as applications to compute the geodesics or level sets in the metric field induced by the diffusion tensors, but do not consider other Riemannian quantities such as curvatures, as is done in this thesis.
3.2 Streamline Tracking of DTI Fibers

Suppose that a white matter body consists only of thick coherent axon bundles as if it was filled with thick electric cables. Then it is likely that in a voxel all fibers have identical orientations. In this case the diffusion tensor has one large eigenvalue and the corresponding eigenvector is aligned with the orientation of fibers. To reconstruct fibers, one would follow the principal eigenvector of the tensors. This amounts to solving a curve \( c(t) \) satisfying the following system

\[
\begin{aligned}
\dot{c} &= \arg \max_{|v|=1} \left( v^T D v \right) \\
c(0) &= p.
\end{aligned}
\] (3.18)

In areas with single dominant fiber population, this tracking method works quite well. For example in [71] streamline reconstructions of white matter tracts in the pons such as those of the middle cerebral peduncle, the medial lemniscus and the superior cerebellar peduncle do indeed agree with the anatomical knowledge of the region. As an example of a relatively high curvature that an axon may have, in Fig. 3.3 we see an illustration of a so-called U-fiber that connects neurons on the cortex to other neurons on the cortex underneath a sulcus, i.e. a "valley".
In a discrete graph with vertices corresponding to spatial points and edges between them encoding the interconnections weighted by mutual distances, we can compute a shortest route between any two vertices, using the algorithm of Dijkstra [25]. This is the discrete counterpart of the variational problem in Sec. 2.5. We take the time of travel to be the distance measure of interest. Just like when traveling by train, we typically talk in terms of travel times rather than distances in kilometers. The real distances i.e. the Euclidean grid around the imaged object is fixed and known. From \( d = v \cdot t \), we see that the greater the mean velocity the shorter the travel time.

Instead of the diffusion tensors we use their inverses \( g = D^{-1} \), i.e. the metric tensors to describe travel time. Once we know the metric tensors we can calculate the edge weights e.g. by taking the mean of the edge weights of adjacent vertices as in Fig. 3.4.

Although the algorithm of Dijkstra will certainly give a shortest path between any two points, it is not likely that bundles with high curvature such
Figure 3.4: Left: Ellipsoids representing the metric tensors on the graph. Right: Distances between adjacent vertices are determined by the metric tensors.

as subcortical bundles will always be the shortest paths. We demonstrate this in Fig. 3.5.

In DTI context, instead of simply using the information of the diffusion tensor, some anatomical knowledge and information about the local neighborhood could be included in the metric tensors. This could result in optimal paths that give better reconstruction of the real anatomical connections. Indeed there are many Hamilton-Jacobi based algorithms, with cost functions derived from the DTI-data, with fast and efficient implementations [51] [47].

Figure 3.5: From left to right: A simulated U-fiber with different widths. The green curves are the shortest paths of Dijkstra. As the angle of opening becomes steeper the shortest path will jump closer to a Euclidean shortest path.
3.4 Quality Measures for DTI Fibers

3.4.1 Connectivity Strength of Pathways

In Sec. 3.2 and 3.3 we saw two prototypes of minimal curves in tensor valued data. Whatever the method of finding these curves, it can be handy to have measures that can rank the curves according to their similarity to real axons.

In [47] a so-called validity index is assigned to each curve that is a numerical solution to a Hamilton-Jacobi system. The validity index

$$\text{VI}(\gamma) = \frac{\int |\dot{\gamma}(s) \cdot \mathbf{v}(s)| ds}{\int ds} = \frac{\int \sqrt{(\delta_{ij}\dot{\gamma}(s)^i\mathbf{v}(s)^j)^2 ds}}{\int \sqrt{\delta_{ij}\dot{\gamma}(s)^i\dot{\gamma}(s)^j} ds}$$

measures how parallel the tangent of a curve is with the principal eigenvector $\mathbf{v}$ of the diffusion tensor along the curve.

We have proposed a measure for curves in [7], which we call the connectivity strength $m(\gamma)$ of a curve $\gamma$ (independently with the definition in [80]). It measures the relative diffusivity along a curve $\gamma$, i.e. the ratio of lengths of $\gamma$ in Euclidean and in diffusion induced Riemannian metric. Let $g = \mathbf{D}^{-1}$ and let $\gamma(t)$ be a parameterized curve and $\dot{\gamma}(t)$ be the unit tangent to the curve. The ratio is

$$m(\gamma) = \frac{\int_0^1 \sqrt{\eta_{ij}(\gamma(t))\dot{\gamma}^i(t)\dot{\gamma}^j(t)} dt}{\int_0^1 \sqrt{g_{kl}(\gamma(t))\dot{\gamma}^k(t)\dot{\gamma}^l(t)} dt},$$

where $\eta_{ij}$ is the covariant Euclidean metric tensor, which in Cartesian coordinates reduces to the constant identity matrix. The connectivity strength is a positive value with higher values indicating better connectivity.

In the neighbourhood of a point $p = \gamma(0)$, the limit of ratio (3.20) gives us a local measure:

$$m(p) = \lim_{t \to 0} \frac{\int_0^t \sqrt{\eta_{ij}(\gamma(t))\dot{\gamma}^i(t)\dot{\gamma}^j(t)} dt}{\int_0^t \sqrt{g_{kl}(\gamma(t))\dot{\gamma}^k(t)\dot{\gamma}^l(t)} dt} = \frac{\sqrt{\eta_{ij}(\gamma(0))\dot{\gamma}^i(0)\dot{\gamma}^j(0)}}{\sqrt{g_{kl}(\gamma(0))\dot{\gamma}^k(0)\dot{\gamma}^l(0)}}.$$
The unit vector that maximizes $m(p)$ is the eigenvector corresponding to the minimal eigenvalue $\lambda_g^-$ of $g$, which in turn corresponds to the maximal eigenvalue $\lambda_D^+$ of $D$ and

$$m(p) = \sqrt{\lambda_g^-} = \frac{1}{\sqrt{\lambda_D^+}}.$$

(3.22)

Thus locally this measure attains maxima in the direction of the principal eigenvector of the DTI-tensor, and agrees with the traditional largest eigenvalue fibre tracking. By splitting up the integrals in (3.20) over a partitioning of the curve $\gamma$ into sub-curves $\gamma_i$, we may apply the measure $m$ also to the sub-curves $\gamma_i$ and measure possible variation in diffusivity along the curve to any desired level of discretization.

Although both measures assign largest values to paths that follow closely the principal eigenvector field, the connectivity strength has some advantages over the validity index. For example, when a curve $\gamma$ extends through a region with almost isotropic diffusion, the principal eigenvector typically becomes unstable and sensitive to noise resulting in an unreliable validity index, whereas the connectivity strength that does not depend on solving eigenvectors, simply records the isotropy by low values.

The diffusivity in the regions of relatively free diffusion (e.g. ventricles) is typically at most two times greater than that in the areas with restricted diffusion (e.g. in thick fiber bundles). While the information on the mean diffusivity is also of importance, if we want to look at the diffusion profile as a probability of fiber orientation, then a reasonable choice is to normalize the diffusion/metric tensors. In view of the magnitude difference in the trace of the Riemann tensor vs. the Euclidean metric tensor in (3.20), normalization makes the interpretation of $m(\gamma)$ easier. When metric tensors are normalized, $m(\gamma) < 1$ ($m(\gamma) > 1$) means that there is in average lower (correspondingly higher) diffusivity along $\gamma$. Another factor that may affect the connectivity measure, is the choice of interpolation method for tensors. As discussed in chapter 6, there are several ways to interpolate tensors. In [5], we experimented with three different interpolation methods, obtaining very similar results w.r.t. connectivity strength measure regardless of the interpolation method chosen.
3.4. Quality Measures for DTI Fibers

Figure 3.6: Left: Geodesic deviation (sectional curvature) measures the relative acceleration of separation of geodesics whose initial direction is a small perturbation of the fiducial reference geodesic (in a plane). Ricci curvature is the mean of geodesic deviation over all planes containing vector $V$. Middle and right: Tubes generated by geodesics that lie in the neighborhood of the fiducial geodesic (blue) and which are initially parallel. The geodesics may stick together, or deviate from each other generating a larger cross sectional area.

3.4.2 Ricci Curvature

We recall the definition of the Ricci curvature

$$ R(v) = \frac{1}{n-1} R_{ik} v^i v^k, \quad (3.23) $$

where $R_{ik}$ is the Ricci tensor (2.33). For an interpretation see the left most picture of Fig. 3.6. A coherent thick bundle of axons will have relatively large number of molecules moving in the direction aligned with this structure. We want to relate this coherence to the Ricci curvature on the diffusion-induced tensor field.

Imagine a ring surrounding a fixed geodesic $\gamma$ and a number of nearby geodesics that are parallel to $\gamma$. We discard all geodesics that are not parallel to these. We then slide the ring along the fixed geodesic for small distance and count how many of the geodesics are still passing through the ring. If the number has decreased the geodesics have deviated from each other, if it has increased the geodesics have merged together (see Fig. 3.6).

We study how the flux of initially parallel nearby geodesics changes in a given direction of interest. For this, we look at the second order rate of change of the flux, which is a relative quantity. A basic observation is that
the flux of these geodesics is inversely proportional to the cross-sectional area occupied by these geodesics.

We consider a rectangular region spanned by two orthogonal (in the Riemannian metric) unit vectors $v_1$ and $v_2$, which are also orthogonal to the tangent of a fixed geodesic at point $t = 0$. Then the vectors $v_1, v_2$ and $\dot{\gamma}$ constitute an orthonormal frame in the tangent space. In Euclidean space, if we translate the vectors $e_1, e_2, e_3$ along a line its Euclidean area will of course not change. Similarly if we parallel transport the vectors $v_1, v_2$ and $\dot{\gamma}$ along a geodesic the Riemannian volume of the frame spanned by these vectors will not change. But in Euclidean sense the volume determined by this triad will change unless the manifold itself is Euclidean. The larger it becomes, the more the geodesics deviate and the smaller the relative flux of geodesics. Since vectors $v_i$ are parallel transported

$$\nabla_{\dot{\gamma}} v_i(0) = 0 , \quad i = 1, 2 . \quad (3.24)$$

Vector fields that are parallel transported along geodesics, satisfy the so-called Jacobi-equation \[19\]

$$\nabla_{\dot{\gamma}} \nabla_{\dot{\gamma}} v^l_s = -R^l_{ijk} \dot{\gamma}^j v^i_s \dot{\gamma}^k . \quad (3.25)$$

From (3.25) we can compute the Taylor approximation for the vector $v_s(t)$ $(s = 1, 2)$ near $\gamma(0)$:

$$v^l_s(t) = v^l_s(0) + \dot{v}^l_s(0) t + \frac{1}{2} \ddot{v}^l_s(0) t^2 + O(t^3)$$

$$= v^l_s(0) - \frac{1}{2} R^l_{ijk} \dot{\gamma}^j(0) v^i_s(0) \dot{\gamma}^k(0) t^2 + O(t^3) . \quad (3.26)$$

Similarly the area $S(t)$ of the quadrilateral spanned by $v_1, v_2$ near $\gamma(0)$:

$$S(t) = S(0) + \dot{S}(0) t + \frac{1}{2} \ddot{S}(0) t^2 + O(t^3) . \quad (3.27)$$

We look at the short time evolution of the plane area defined by the vectors $v_i$ and their parallel transports. Let $\mathbf{n}$ be a unit vector orthogonal to $v_1, v_2$.

$$\mathbf{n} = \frac{v_1 \times v_2}{|v_1 \times v_2|} . \quad (3.28)$$
The plane area spanned by vectors $v_1$ and $v_2$ is then as follows (abbreviating $v_i := v_i(0)$ on the right hand side).

\[
S(0) = ((v_1 \times v_2) \cdot (v_1 \times v_2))^{\frac{1}{2}}.
\]
\[
\dot{S}(0) = ((v_1 \times v_2) \cdot (v_1 \times v_2))^{-\frac{1}{2}} (v_1 \times v_2) (\ddot{v}_1 \times v_2 + v_1 \times \ddot{v}_2)
= 0
\]
\[
\ddot{S}(0) = ((v_1 \times v_2) \cdot (v_1 \times v_2))^{-\frac{1}{2}} (v_1 \times v_2) \cdot (\dddot{v}_1 \times v_2 + v_1 \times \dddot{v}_2)
= n \cdot (\dddot{v}_1 \times v_2 + v_1 \times \dddot{v}_2)
= -n \cdot (v_2 \times \dddot{v}_1) - n \cdot (v_1 \times \dddot{v}_2)
= v_1 \dddot{v}_1 + v_2 \dddot{v}_2
= \frac{1}{2} \left( R_{ijk} \dot{\gamma}^i \gamma^j v_1^l \dot{\gamma}_k v_1^l + R_{ijk} \dot{\gamma}^i \gamma^j v_2^l \dot{\gamma}_k v_2^l \right)
= R_{ik} \dddot{\gamma}^i \dot{\gamma}^k.
\]

The approximate limiting area of the quadrilateral patch generated by the geodesics (that penetrate the initial quadrilateral orthogonally) is thus

\[
S(t) = S(0) - \frac{1}{2} R_{ik}(\gamma(0)) \dot{\gamma}^i(0) \dot{\gamma}^k(0) t^2 + O(t^3).
\]

Non-negative Ricci curvature implies stability or increase in flux and relatively stable bundle of geodesics, while negative Ricci curvature implies decrease in flux and less coherent bundle of geodesics, which might imply bifurcation or crossing of the fiber bundle. For an illustration of Ricci curvature computed on tracked streamlines see Fig. 3.7. We remark that the standard DTI streamlines generally cannot propagate through a crossing as can be seen also in this example.

### 3.4.3 Stickiness

Now that we have an interpretation of the Ricci curvature, we may ask: In what direction does the Ricci curvature attain a maximum? Since the Ricci tensor is symmetric, this question has an easy answer: in the direction of the greatest positive eigenvector (although it may not exist at every point). If the tensor field is homogeneous, the Ricci tensor is zero, giving no information of a preferred (in sense of coherence) direction. To assess the degree that the locally optimal direction (principal eigenvector of the diffusion tensor) is in coherence with the neighboring optima one
can compute the inner product of the Ricci principal eigenvector and principal eigenvector of the diffusion tensor. This, which we call the pointwise stickiness of a curve, gives information about the credibility of the local structure to belong to a thick/stable bundle.

In Fig. 3.8 we illustrate the stickiness measure on a crossing.

### 3.4.4 Inhomogeneity Detection

In Riemannian geometry, it is a well known fact that the space is homogeneous i.e. the metric tensors are identical if the Ricci tensor is proportional to the metric tensor. Then the space will have constant curvature. Since the Ricci tensor assigns to each spatial position and a vector (direction) the second order rate of change of Euclidean distance of initially parallel...
3.5 Ricci Scalar

Different scalar measures have been proposed in the literature, aiming to reveal first hand information about the underlying diffusion process and
tissue structure. Scalar measures are also studied to relate the state of white matter to different pathologies. The most common ones are the mean diffusivity (MD)

\[ \frac{1}{3} (\lambda_1 + \lambda_2 + \lambda_3) \]  

(3.30)

where \( \lambda_i \) is the \( i \)th eigenvalue of diffusion tensor \( \mathbf{D} \) and the fractional anisotropy (FA)[12]

\[ \frac{1}{\sqrt{2}} \frac{\sqrt{(\lambda_1 - \lambda_2)^2 + (\lambda_1 - \lambda_3)^2 + (\lambda_2 - \lambda_3)^2}}{\sqrt{\lambda_1^2 + \lambda_2^2 + \lambda_3^2}} \]  

(3.31)

Mean diffusivity can be interpreted as average diffusion per voxel. In general, MD is smaller in areas with organized tissue compared to e.g. ventricles which are filled with cerebrospinal fluid. Fractional anisotropy measures the anisotropy of the diffusion.
For convenience we recall here the Definition 2.41 of the Ricci scalar.

\[ R = R_{ik} g^{ik}. \]  

(3.32)

It is a trace of the Ricci tensor.

A zero FA value indicates the lack of dominant direction of diffusion (see Fig. 3.10 left) while a non-zero value indicates anisotropy to some degree, i.e., the presence of structures obstructing the diffusion process [?]. On the other hand, a vanishing Ricci scalar can indicate both isotropy or anisotropy (see Fig. 3.10 first two from the left). The same is true for non-zero Ricci scalars (see Fig. 3.10 last two on the right). In the literature of applied differential geometry [84] the Ricci scalar has been used for curvature analysis of 2D medical images.

3.5.1 Experimental Results

We have experimented by computing Ricci scalars on simulated, phantom and real data. The computations were done using Mathematica 7.0. For the derivatives on a discrete data volume, we used the Gaussian derivatives implemented in a locally designed Mathematica package [94].

Simulated data

To get a quick insight in what the Ricci scalar can detect in a tensor field, we refer to Fig. 3.11, where we have simulated a crossing of orthogonally
oriented sets of tensors, modeling diffusion tensors corresponding to two fiber bundles. The voxels where the Ricci scalar is non-zero are colored according to its sign using a temperature map. In the crossing region of this tensor field, the Ricci scalar tends to be large and negative. Since the Ricci scalar involves second order derivatives, we can tune the minimum size of the region to be considered by manipulating the scale of the Gaussian differential operator [30] [42] [35].

![Figure 3.11: Top left: A simulated crossing of fibers. Top right: A slice of a tensor field simulating the crossing. Bottom left: A horizontal slice on the boundary of a fiber, blue color corresponding to negative Ricci scalars. Bottom right: A horizontal slice in the middle of the crossing with slightly negative Ricci scalars in the crossing.](image)

**Phantom data**

We computed Ricci scalars on a real phantom data consisting of cylinder containing a water solution, three sets of crossing synthetic fiber bundles and three supporting pillars on the boundary. In Fig. 3.12 we see that
in the region where the fiber bundles cross Ricci scalars have relatively large negative values, despite of the noisy nature of the DTI-data. The DT-image of the phantom has dimensions $10 \times 104 \times 104$ and therefore the resolution in vertical direction is far less than in horizontal directions.

![Image of phantom and DTI data](image)

Figure 3.12: Left: A phantom (courtesy of Pim Pullens, Maastricht University) containing intersecting bundles of synthetic fibers. Middle: A temperature map of Ricci scalars on a horizontal slice. Right: Ricci scalars on a vertical slice containing a crossing. Large negative values are found in the crossing.

**Real data**

We have also experimented with real DTI data of a rat brain. We plotted the Ricci scalars in a temperature map, to emphasize the differences in sign. We identified positive (negative) outliers of the Ricci scalar data with maximum (minimum) values of the rest of the data. The Ricci scalar gives information about the variations in diffusion tensor orientations unlike FA, which will identify tensors with similar anisotropy even though their orientation may differ. This can be seen e.g. in the region surrounded by a pink square in Fig. 3.13, which is known to have complex structure and contain two larger bundles of fibers (internal- and external capsules) with different orientation [15].
Figure 3.13: Left: Ricci scalars on a slice of the rat brain DTI image. Middle: Fractional anisotropy. Right: Mean diffusivity.
"A. Einstein used Riemannian geometry to describe his general relativity theory, assuming that spacetime is always Riemannian." Zhongmin Shen
4.1 Introduction

According to S.S. Chern, Finsler geometry is just Riemannian geometry without the quadratic restriction. Although not as popular as Riemann geometry, Finsler geometry is also widely studied and applied successfully in theoretical physics etc. [2][22][1][89]. Modern books on Finsler geometry began to appear at increasing rate in late 1990’s [10][86][68] while some older books have paved the way [27][18][82]. Already in classical books on calculus of variations, a homogeneous Lagrangian is called Finsler arc length [100]. The idea of Finsler geometry was already in the considerations of Riemann, to have a norm function that depends homogeneously on a line-element in addition to position. In this chapter we begin with the basic definitions, and provide a new result concerning strong convexity. We define the connection as in [10], and scrutinize the derivation of the equation for geodesics by [86].

Let $M$ be a $C^\infty$ manifold, $T_xM$ the tangent space at each $x$ and $TM = \{(x,y) \mid x \in M, y \in T_xM\}$.

**Definition 4.1.** A function $F : T_xM \to [0, \infty)$ is called a Minkowski norm if $F$ satisfies the following conditions:

(M1) $F(\lambda y) = \lambda F(y)$ for every $\lambda > 0$ and $y \in T_xM$.

(M2) $F$ is $C^\infty$ on $T_xM \setminus \{0\}$.

(M3) The bilinear form $g_{ij}(y) = \frac{1}{2} \frac{\partial^2 F^2}{\partial y_i \partial y_j}$ in $T_xM$ is positive definite.

A function $F : TM \to [0, \infty)$ is called a Finsler metric if it has the following properties:

1. $F$ is $C^\infty$ on $TM \setminus \{0\}$.

2. for each $x \in M$, $F|_{T_xM}$ is a Minkowski norm on $T_xM$.

This bi-linear form

$$g_{ij}(x,y) = \frac{1}{2} \frac{\partial^2 F^2}{\partial y^i \partial y^j}, \quad (4.1)$$

is called the Finsler metric tensor. Note that if $F(x,y) = \sqrt{g_{ij}(x)y^i y^j}$, then $g_{ij}(x,y) = \frac{1}{2} \frac{\partial^2}{\partial y^i \partial y^j} (g_{ij}(x)y^i y^j) = g_{ij}(x)$ is the usual Riemannian metric.
4.2 Homogeneity of Finsler Norm

We recall here some consequences of the homogeneity (M1) of the norm function, since these are useful in the sequel. Let \( x \in M \) and \( y \in T_x M \), and \( F(x, y) \) be a Minkowski norm on \( TM \setminus \{0\} \). We denote the partial derivative of a scalar-valued function \( F(x, y) \) as

\[
\frac{\partial F}{\partial y^i} := F_{y^i} . \tag{4.2}
\]

From the homogeneity of \( F \)

\[
F(x, \lambda y) = \lambda F(x, y) . \tag{4.3}
\]

Taking derivatives w.r.t. \( \lambda \)

\[
F_{y^i} y^i = F , \tag{4.4}
\]

and w.r.t. \( y^j \)

\[
F_{y^i} y^i y^j = 0 . \tag{4.5}
\]

4.3 Strong convexity of Finsler Norm

4.3.1 Strong convexity criterion

We derive another criterion for strong convexity in dimension three, in analogy to the corresponding criterion in dimension two [10].

**Theorem 4.2.** Let us consider the inequality

\[
g_{ij}(y)v^i v^j > 0 , \tag{4.6}
\]

where \( y \in T_x M \) and \( v \in \mathbb{R}^3 \). From the homogeneity of the norm function \( F \), it follows that it is sufficient to have this condition on the unit level set of the norm. We consider this level surface i.e. the so-called indicatrix, which is the set

\[
\{ u \in T_x M \mid F(u) = 1 \} . \tag{4.7}
\]

and a parametrization

\[
u(\theta, \varphi) = (u^1(\theta, \varphi), u^2(\theta, \varphi), u^3(\theta, \varphi)) . \tag{4.8}\]
We define the following three matrices:

\[ m = \begin{pmatrix} u^1 & u^2 & u^3 \\ u_\theta^1 & u_\theta^2 & u_\theta^3 \\ u_\phi^1 & u_\phi^2 & u_\phi^3 \end{pmatrix}, \quad m_\theta = \begin{pmatrix} u^1_\theta & u^2_\theta & u^3_\theta \\ u^1_\phi & u^2_\phi & u^3_\phi \\ u^1_\phi & u^2_\phi & u^3_\phi \end{pmatrix}, \quad m_\phi = \begin{pmatrix} u^1_\phi & u^2_\phi & u^3_\phi \\ u^1_\phi & u^2_\phi & u^3_\phi \\ u^1_\phi & u^2_\phi & u^3_\phi \end{pmatrix}. \] 

(4.9)

Then the inequality (4.6) is satisfied if and only if the following inequalities are valid.

\[ \frac{\det(m_\theta)}{\det(m)} < 0, \] 

(4.10)

\[ \frac{\det(m_\phi)}{\det(m)} < (g_{ij}u^i_\theta u^j_\phi)^2 \frac{\det(m)}{\det(m_\theta)}. \] 

(4.11)

Proof.

In what follows we abbreviate \( g_{ij} = g_{ij}(x, u) \). From \( F(u) = 1 \) we have

\[ g_{ij}u^i u^j = 1. \] 

(4.12)

Taking derivatives of both sides and using a consequence of Euler’s theorem for homogeneous functions ([10] p.23) that says

\[ \frac{\partial g_{ij}}{\partial u^k} u^k = 0, \] 

(4.13)

we obtain

\[ g_{ij}u^i_\theta u^j_\theta = 0, \] 

\[ g_{ij}u^i_\phi u^j_\phi = 0, \] 

(4.14)

implying \( u_\theta \perp_g u \) and \( u_\phi \perp_g u \), since \( u_\theta \neq 0 \) and \( u_\phi \neq 0 \) (except at \( \theta = 0 \)).

Taking derivatives once more, we get

\[ g_{ij}u^i_\theta u^j_\theta = -g_{ij}u^i_\phi u^j_\phi, \] 

\[ g_{ij}u^i_\phi u^j_\phi = -g_{ij}u^i_\phi u^j_\phi, \] 

\[ g_{ij}u^i_\theta u^j_\phi = -g_{ij}u^i_\phi u^j_\phi. \] 

(4.15)
We may express an arbitrary vector \( \mathbf{w} \) as a linear combination of orthogonal basis vectors:

\[
\mathbf{w} = \alpha \mathbf{u} + \beta \mathbf{u}_\theta + \gamma \left( \mathbf{u}_\varphi - \frac{\langle \mathbf{u}_\varphi, \mathbf{u}_\theta \rangle}{\langle \mathbf{u}_\theta, \mathbf{u}_\theta \rangle} \mathbf{u}_\theta \right).
\]  

(4.16)

We substitute this expression for \( \mathbf{w} \) to the left hand side of (4.6) and obtain:

\[
g_{ij}w^i w^j = \alpha^2 g_{ij} u^i u^j - \beta^2 g_{ij} u^i_{\theta\theta} u^j - \gamma^2 \left( g_{ij} u^i_{\varphi\varphi} u^j + \frac{(g_{ij} u^i_{\theta} u^j_{\varphi})^2}{g_{ij} u^i_{\theta} u^j_{\theta}} \right),
\]

because the mixed terms vanish due to the orthogonality of basis vectors.

On the other hand, for \( \mathbf{u} \)'s on the indicatrix we have as a consequence of Euler's theorem on homogeneous functions (denoting \( F_{\mathbf{u}^i} = \frac{\partial F}{\partial \mathbf{u}^i} \)):

\[
F_{\mathbf{u}^i} \mathbf{u}^i = F(\mathbf{u}) = 1 .
\]

(4.18)

Differentiating (4.18) w.r.t. \( \theta \) and \( \varphi \), we obtain two equations:

\[
F_{\mathbf{u}^i} u^i_{\theta} = 0 ,
\]

(4.19)

\[
F_{\mathbf{u}^i} u^i_{\varphi} = 0 ,
\]

(4.20)

for \( F \) is a homogeneous function.

The matrices \( m, m_\theta, m_\varphi \) are as defined in (4.9). Solving system of equations (4.18), (4.19) and (4.19) we get:

\[
F_{\mathbf{u}^i} = -\frac{u^2_{\varphi} u^3_{\theta} - u^2_{\theta} u^3_{\varphi}}{\det(m)},
F_{\mathbf{u}^2} = -\frac{u^3_{\varphi} u^1_{\theta} - u^3_{\theta} u^1_{\varphi}}{\det(m)},
F_{\mathbf{u}^3} = -\frac{u^1_{\varphi} u^2_{\theta} - u^2_{\varphi} u^1_{\theta}}{\det(m)} .
\]

(4.21)

Now using equalities

\[
F_{\mathbf{u}^i} g_{ij} u^j, \quad g_{ij} u^i_{\theta\theta} u^j = F_{\mathbf{u}^k} u^k_{\theta},
\]

\[
g_{ij} u^i_{\varphi\varphi} u^j = F_{\mathbf{u}^k} u^k_{\varphi\varphi} ,
\]

(4.22)
and
\[ g_{ij} u^i_{\theta} u^j = \frac{\det(m_{\theta})}{\det(m)} , \quad g_{ij} u^i_{\varphi} u^j = \frac{\det(m_{\varphi})}{\det(m)} , \] (4.23)
we obtain
\[ g_{ij} w^i w^j = \alpha^2 - \beta^2 g_{ij} u^i_{\theta} u^j \\
- \gamma^2 \left( g_{ij} u^i_{\varphi} u^j + \left( g_{ij} u^i_{\theta} u^j \right)^2 \right) > 0 \] (4.24)
iff
\[ \frac{\det(m_{\theta})}{\det(m)} < 0 \quad \text{and} \quad \frac{\det(m_{\varphi})}{\det(m)} < \left( g_{ij} u^i_{\theta} u^j \right)^2 \frac{\det(m)}{\det(m_{\theta})} . \] (4.25)

Note that the right hand side of (4.11) is definitely non-positive. This result implies geometrically that the systems of three vectors in matrices \( m \) have opposite orientation relative to those in \( m_{\theta} \) and \( m_{\varphi} \). \( \square \)

For illustration see Fig. 4.1.

### 4.3.2 Triangle inequality

We show that strong convexity implies triangle inequality. Let us thus assume that (suppressing \( y \) from notation, i.e. \( g_{ij} := g_{ij}(y) \))
\[ g_{ij} v^i v^j = \frac{1}{2} \frac{\partial^2 F^2}{\partial y^i \partial y^j} v^i v^j > 0 , \forall v \in \mathbb{R}^n . \] (4.26)
Then \( g_{ij} \) defines an inner product, and the Cauchy-Schwarz inequality says that
\[ \left( g_{ij} \eta^i \xi^j \right)^2 \leq \left( g_{ij} \eta^i \eta^j \right) \left( g_{ij} \xi^i \xi^j \right) , \forall \eta, \xi \in \mathbb{R}^n . \] (4.27)
By choosing \( \eta = y \) we obtain the following equation
\[ \left( g_{ij} y^i \xi^j \right)^2 \leq F^2(y) \left( g_{ij} \xi^i \xi^j \right) , \forall \xi \in \mathbb{R}^n . \] (4.28)
Since
\[ g_{ij} \xi^i \xi^j = F y^i y^j \xi^i \xi^j + F^i y^j , \] (4.29)
it follows from (4.28) that
\[
F_{y^i y^j} \xi^i \xi^j = \frac{1}{F} \left( g_{ij} \xi^i \xi^j - F_{y^i} F_{y^j} \xi^i \xi^j \right) \\
= \frac{1}{F^3} \left( F^2(y) g_{ij} \xi^i \xi^j - FF_{y^k} \xi^k FF_{y^l} \xi^l \right) \\
= \frac{1}{F^3} \left( F^2(y) g_{ij} \xi^i \xi^j - (g_{kl} y^k \xi^l)^2 \right) \\
\geq 0 ,
\]
where equality holds only in case \( \xi = \lambda y \) for some \( \lambda \in \mathbb{R} \setminus \{0\} \). Next, it is shown that
\[
2F(y) \leq F(y + \xi) + F(y - \xi) .
\]
1. Suppose that $\xi = \lambda y$. If $|\lambda| \leq 1$, then
\[ F((1+\lambda)y) + F((1-\lambda)y) = (1+\lambda)F(y) + (1-\lambda)F = 2F(y) \, . \quad (4.32) \]
If $|\lambda| > 1$, then
\[ F((y+\lambda)y) + F((y-\lambda)y) = (2+\alpha)F(y) + \beta \, , \quad \alpha, \beta > 0 \, . \quad (4.33) \]
If $\xi = 0$, $F(y) + F(y) \geq 2F(y)$.

2. Suppose that $\xi \neq \lambda y$. The second mean value theorem states that there exist a $0 < \delta < 1$, such that
\[ F(y + \pm \xi) = F(y) \pm F_{y^i} \xi^i + \frac{1}{2}F_{y^i y^j}(y \pm \varepsilon \xi) \xi^i \xi^j \, , \quad (4.34) \]
for some $0 < \varepsilon < 1$. From (4.30) we obtain finally that
\[ F(y + \xi) + F(y - \xi) > 2F(y) \, , \quad (4.35) \]
which gives the usual triangle-inequality by substitutions $y = \frac{1}{2}(y_1 + y_2)$ and $\xi = \frac{1}{2}(y_1 - y_2)$.

### 4.4 Connection

In Riemann spaces (Chapter 2), it was shown that metric compatibility (2.5) and torsion freeness (2.4) resulted in a unique linear connection on the manifold. In classical theoretical physics [67], the equivalence principle states that "the laws of physics are the same in any local Lorentz coordinate frame of curved spacetime, as in a global one of a flat spacetime." According to [67], a necessary condition for this principle to hold is that for vectors $v, u, w$ for which $v + u = w$ at point $\gamma(0)$, this relation should also hold after parallel transporting them along a geodesic $\gamma$ to a point $\gamma(t)$. This in turn implies that there is no torsion (2.7). Similarly, torsion freeness can be taken as a prerequisite for a Finsler connection. A connection is said to be almost metric compatible if the connection (or covariant derivative) satisfies the product rule up to a term given by the so called Cartan tensor ([10], p.38). Torsion freeness and almost metric compatibility define a unique linear connection on the pulled-back bundle $\pi^*TM$, called the Chern-connection. Although the details are beyond the scope of this thesis and we have not here implemented nor applied the Chern
connection, we mention the interesting fact that a Finsler connection cannot be both torsion free and satisfy the product rule as the Levi-Civita connection in Riemann space. An alternative to Chern connection is the Cartan connection which satisfies the product rule but is not torsion free.

A connection gives us the means to take directional derivatives. If the connection coefficients agree to the structural equations (torsion freeness and almost metric compatibility), we can replace \( \frac{\partial}{\partial x^i} \) by

\[
\frac{\delta}{\delta x^i} := \frac{\partial}{\partial x^i} - N^j_i \frac{\partial}{\partial y^j},
\]

(4.36)

where

\[
N^j_i := \gamma^j_{ik} y^k - C^j_{ik} \gamma^k_r y^r y^s.
\]

(4.37)

And further

\[
\gamma^i_{jk} := \frac{g^{is}}{2} \left( \frac{\partial g_{sj}}{\partial x^k} + \frac{\partial g_{sk}}{\partial x^j} - \frac{\partial g_{kj}}{\partial x^s} \right),
\]

(4.38)

and

\[
C^i_{jk} := g^{is} C_{jsk},
\]

(4.39)

with

\[
C_{ijk} := \frac{1}{4} \frac{\partial^3 F^2}{\partial y^i \partial y^j \partial y^k}.
\]

(4.40)

Note that the Chern connection agrees with the Levi-Civita connection on a Riemannian space where \( C_{ijk} = 0 \).

## 4.5 Geodesics

We scrutinize the derivation of equations for geodesics following the presentation in [86]. Let \( \gamma : [a, b] \to M \) be a constant speed piecewise \( C^\infty \) curve

\[
F(\gamma, \dot{\gamma}) = \lambda,
\]

(4.41)

with \( \lambda > 0 \). Then there is a partition of \([a, b] : a = t_0 < \cdots < t_k = b\), so that on every subinterval \([t_{i-1}, t_i], \gamma \) is \( C^\infty \). We fix this partition and define a piecewise \( C^\infty \) map \( H : (\varepsilon, \varepsilon) \times [a, b] \to M \) (a variation of \( \gamma \)) to be such that:
• $H$ is $C^0$ on $(\varepsilon, \varepsilon) \times [a, b]$

• $H$ is $C^\infty$ on each $(\varepsilon, \varepsilon) \times [t_{i-1}, t_i]$, $i = 0, \ldots, k$.

• $\gamma(t) = H(0, t), \quad a \leq t \leq b$.

We also define a vector field $V(t)$, called the variation field of $H(u, t)$ as follows:

$$V(t) = V^i(t) \frac{\partial}{\partial x^i} |_{\gamma(t)} := \frac{\partial H}{\partial u}(0, t). \quad (4.42)$$

Then the length of $\gamma_u(t) := H(u, t)$ is given by

$$L(u) = \int_a^b F(\gamma_u(t), \dot{\gamma}(t)) \, dt = \sum_{i=1}^k \int_{t_{i-1}}^{t_i} F\left(\gamma_u(t), \frac{\partial H}{\partial t}(u, t)\right) \, dt. \quad (4.43)$$

We want to find out the extremal point of this variational curve, i.e. the condition when $\frac{\partial L}{\partial u}(0) = 0$.

First we work out an expression for $\frac{\partial L}{\partial u}(0) := L'$. Except for the first few terms, we suppress the arguments of $F$ for brevity.

$$L'(0) = \int_a^b \frac{\partial}{\partial u} (F(\gamma_u(t), \dot{\gamma}_u(t))) |_{u=0} \, dt$$

$$= \int_a^b \frac{\partial}{\partial u} \left( F(\gamma(t), \frac{\partial H}{\partial t}(u, t))\right) |_{u=0} \, dt$$

$$= \int_a^b \left( \frac{\partial F}{\partial x^k} \frac{\partial^2 \gamma_k}{\partial t^2} + \frac{\partial F}{\partial y^k} \frac{\partial \gamma^k}{\partial u}\right) |_{u=0} \, dt$$

$$= \int_a^b \left( \frac{\partial F}{\partial x^k} \frac{\partial H^k}{\partial u} + \frac{\partial F}{\partial y^k} \frac{\partial^2 H}{\partial u \partial t}(u, t)\right) |_{u=0} \, dt$$

$$= \int_a^b \left( \frac{\partial F}{\partial x^i} V^i(t) + \frac{\partial F}{\partial y^i} \frac{dV^i}{dt}(t)\right) \, dt. \quad (4.44)$$

Secondly, since $F > 0$ (4.41) we have the following

$$\frac{\partial F}{\partial x^k} V^k + \frac{\partial F}{\partial y^k} \frac{dV^k}{dt} = \frac{1}{2F} \left( \frac{\partial F^2}{\partial x^k} V^k + \frac{\partial F^2}{\partial y^k} \frac{dV^k}{dt}\right). \quad (4.45)$$
In other words

\[ L'(0) = \int_a^b \frac{1}{2F} \left( \frac{\partial F^2}{\partial x^k} V^k + \frac{\partial F^2}{\partial y^k} \frac{dV^k}{dt} \right) dt. \] (4.46)

This can be rewritten as follows

\[ L'(0) = \int_a^b \frac{1}{2F} \left( \frac{\partial F^2}{\partial x^k} V^k - \frac{d}{dt} \left( \frac{1}{2F} \frac{\partial F^2}{\partial y^k} \right) V^k dt + \sum_{i=1}^k \left[ \frac{1}{2F} \frac{\partial F^2}{\partial y^k} V^k \right]_{t_i-1} \]

\[ = \int_a^b \frac{1}{2F} \left( \frac{\partial F^2}{\partial x^k} - \frac{\partial^2 F^2}{\partial y^k \partial x^l} \dot{x}^l - \frac{\partial^2 F^2}{\partial y^k \partial y^l} \dot{y}^l \right) V^k dt \]

\[ + \sum_{i=1}^k \left[ \frac{1}{2F} \frac{\partial F^2}{\partial y^k} V^k \right]_{t_i-1} \]

\[ = \int_a^b \frac{1}{2F} \left( \frac{\partial F^2}{\partial x^k} - \frac{\partial^2 F^2}{\partial y^k \partial x^l} \dot{\gamma}^l - \frac{\partial^2 F^2}{\partial y^k \partial y^l} \ddot{\gamma}^l \right) V^k dt \]

\[ + \left[ \frac{1}{F} g_{jk} \dot{\gamma}^j V^k \right]_{t_i-1}. \] (4.47)

By defining

\[ G^i(y) = \frac{1}{4} g^{il}(y) \left( \frac{\partial^2 F^2}{\partial y^l \partial x^k} y^k - \frac{\partial F^2}{\partial x^l} \right), \] (4.48)

we see that

\[ \frac{2}{F} g_{jk} G^j = \frac{1}{2F} \left( \frac{\partial F^2}{\partial x^l} \frac{y^l}{y^k} - \frac{\partial F^2}{\partial x^k} \right). \] (4.49)

We have finally

\[ L'(0) = -\int_a^b \frac{1}{F} g_{ik} \left\{ 2G^i(\dot{\gamma}) + \ddot{\gamma}^i \right\} V^k dt + \sum_{i=1}^k \left[ \frac{1}{F} \dot{\gamma}^i g_{ik}(\dot{\gamma}) V^k \right]_{t_i-1}. \] (4.50)

**Definition 4.3.** The so-called geodesic curvature \( \kappa(t) \) is defined as follows:

\[ \kappa(t) := \frac{1}{F(\dot{\gamma}(t))^2} \left( \ddot{\gamma}^i + 2G^i(\dot{\gamma}) \right) \frac{\partial}{\partial x^i} |_{\gamma(t)}. \] (4.51)
Denoting with $g_{\hat{\gamma}(t)}\langle \cdot, \cdot \rangle$ the fundamental bi-linear form, we can write (4.50) in an index-free expression:

$$L'(0) = -\lambda \int_a^b g_{\hat{\gamma}(t)}\langle \kappa, V \rangle dt + \frac{1}{\lambda} \left( g_{\hat{\gamma}(b)}\langle \hat{\gamma}(b), V(b) \rangle - g_{\hat{\gamma}(a)}\langle \hat{\gamma}(a), V(a) \rangle \right)$$

$$+ \frac{1}{\lambda} \sum_{i=1}^{k-1} \left( g_{\hat{\gamma}(t_i^-)}\langle \hat{\gamma}(t_i^-), V(t_i) \rangle - g_{\hat{\gamma}(t_i^+)}\langle \hat{\gamma}(t_i^+), V(t_i) \rangle \right).$$

(4.52)

For an arbitrary piecewise variation, that fixes the variation curve to coincide with the original constant speed curve $\gamma$ at points $t_i$, i.e. $H(u, t_i) = \gamma(t_i)$, we have $V(t_i) = 0$ and Eq. (4.52) becomes:

$$L'(0) = -\frac{1}{\lambda} \int_a^b g_{\hat{\gamma}(t)}\langle \kappa, V \rangle dt = 0 ,$$

(4.53)

for a curve $\gamma(t)$ with minimal length. This implies that $\kappa = 0$ i.e. that

$$\dddot{\gamma}^i + 2G^i(\dot{\gamma}) = 0 .$$

(4.54)

This proves that, if $\gamma$ is a piecewise, constant speed, $C^\infty$ curve in Finsler space, and has a minimal length then its geodesic curvature $\kappa = 0$. 

Applications in HARDI

"Tien regels voor succes. . . . Maar, beste lezers, al deze goed bedoelde adviezen zijn slechts een aanloop tot mijn tiende en laatste gouden regel die alle andere overstijgt: Luister niet naar advies." Robbert Dijkgraaf
5.1 Introduction

High angular resolution diffusion imaging (HARDI) is a collective name for techniques [103][91][52] that acquire more directional measurements improving the angular resolution and incorporate a more complex model of local diffusion than the rank two DTI tensor that we discussed in Chapter 3. A HARDI scan acquires diffusion weighted images, just like the Diffusion Tensor Imaging (DTI), but typically with higher angular frequency ranging from 50 to 200 directions per voxel. The so called apparent diffusion coefficient (ADC) \( D(y) \), is again computed from the Stejskal-Tanner [88] formula

\[
\frac{S(y)}{S_0} = \exp(-bD(y)),
\]

where \( S(y) \) is the signal associated with gradient direction \( y \), \( S_0 \) the signal obtained when no diffusion gradient is applied, the b-value as in (3.5) in Chapter 3. It has been shown experimentally [95], that the local maxima of apparent diffusion coefficients, computed using Eq. (5.1) do not correspond to the maxima of actual diffusion when a voxel contains multiple fiber populations, that have distinct orientations. To overcome this problem, several methods have been proposed. For an overview on these methods we refer to [23]. From validations using phantoms [78] and simulations [102][52][104], we learn however that when three or more fiber orientations are present, they are at best recovered only up to 10–30 % of accuracy.

From the various HARDI models we adopt here the so-called Q-ball imaging technique [92][24] and model HARDI measurements using spherical tensors, i.e. homogeneous polynomials restricted to the sphere. Our method is by no means restricted to Q-ball imaging, but this makes regularization, which is a subject of Chapter 6 particularly easy.

We recall that in the Diffusion Tensor Imaging framework, Eq. (5.1) is interpreted as

\[
\frac{S(y)}{S_0} = \exp(-by^T Dy),
\]

with the \( 3 \times 3 \) two-tensor \( D \) describing the probability of directional diffusivity at each voxel. In chapter 3, we used the inverse of the diffusion tensor \( D \) as the Riemann metric tensor. This approach has been exploited to some extent in the DTI literature [75],[60],[7],[5].

Here we study HARDI images as metric spaces, as in the DTI case (3), but
using Finsler metric as a more descriptive model for directional information than the local position dependent inner product i.e. Riemannian metric. Using a tensor representation, one can construct a Finsler norm suitable for the analysis of diffusion profiles that can have a more general shape than that of an ellipsoid.

One of the main results in this Chapter is to show how the ODF can be directly computed from a single tensor fitted to data. This is a shortcut compared to iteratively fitting different order harmonics, the knowledge of which is essential for constructing ODF. As another result we introduce a novel method for fiber tracking in HARDI data based on Finsler geometry. Finsler geometry has been introduced in HARDI setting already in [66], where a dynamic programming approach is taken to compute optimal curves w.r.t. a Finsler norm, without computing the Finsler metric tensors and the accompanying eigenvectors as we do here for more detailed information. In contrast to their work, we use a tensor model, ODE-based tracking and a different homogenization technique.

We recall that in a perfectly homogeneous and isotropic medium, geometry is Euclidean, and shortest paths are straight lines. In an inhomogeneous space, where the metric tensor depends on a position, geometry is Riemannian and the shortest paths are geodesics induced by the Levi-Civita connection [19]. If a medium is not only inhomogeneous, but also anisotropic, i.e. has innate directional structure, the appropriate geometry is Finslerian [10][86] and the shortest paths are correspondingly Finsler-geodesics. As a consequence of the anisotropy, the metric tensor (cf. Eq. (5.11)) depends on both position and direction. This is also a natural model for high angular resolution diffusion images.

5.2 Approximating Data with Tensors

In HARDI literature spherical harmonics are a popular choice to represent functions on the sphere [36][43][24][4]. In this paper we prefer an equivalent tensor (a polynomial consisting of monomials of fixed degree) representation [101][11][31][32] to utilize tensor based differential geometry.

We consider the diffusion profile at a single radial shell and therefore the input vectors to the tensors are unit vectors. This results in a unique
symmetric tensor representation of data that is intuitively a direct generalization of the diffusion tensor with a similar symmetric array notation. A symmetric tensor here means that

\[ D_{i_1 \cdots i_n} = D_{\sigma(i_1) \cdots \sigma(i_n)} , \]  

(5.3)

for any permutation \( \sigma \). Indeed also the data fitting procedure is a straightforward extension of that described by equations (3.14)-(3.17) in Chapter 3. The only difference is that higher order symmetric tensors have more distinct monomials in vector (3.15). The number \( N \) of distinct monomials of order \( k \) chosen from \( n \) variables solves an elementary multi-choose problem [46] and is

\[ N = C(n + k - 1, n) = \frac{(n + k - 1)!}{(k - 1)!n!} . \]  

(5.4)

In case of 4th order tensor we have thus 15 distinct monomials, for 6th order this is 28 etc.

Although in Chapter 2 we have defined the metric tensor and various curvature tensors in a different setting, for convenience we recall briefly the definition of a general tensor.

A (covariant) tensor \( D \) of degree \( n \) at a point \( p \) on a differentiable manifold \( M \) can be seen as a multilinear mapping

\[ D(p) : T_pM \times \cdots \times T_pM \rightarrow \mathbb{R} , \]  

(5.5)

in which \( T_pM \) denotes the tangent space at \( p \in M[56] \). For example, a second order symmetric spherical tensor in \( \mathbb{R}^3 \) can be written as

\[
D_{ij}y^iy^j = D_{11}y^1y^1 + D_{12}y^1y^2 + D_{13}y^1y^3 \\
+ D_{21}y^2y^1 + D_{22}y^2y^2 + D_{23}y^2y^3 \\
+ D_{31}y^3y^1 + D_{32}y^3y^2 + D_{33}y^3y^3 \\
= D_{11}y^1y^1 + 2D_{12}y^1y^2 + 2D_{13}y^1y^3 \\
+ D_{22}y^2y^2 + 2D_{23}y^2y^3 + D_{33}y^3y^3 
\]

(5.6)

suppressing the fiducial point \( p \) in the notation and defining

\[ y = (y^1, y^2, y^3) \]

= \( (\sin \theta \cos \varphi, \sin \theta \sin \varphi, \cos \theta) \).

(5.7)
To justify the choice of tensors we recall that in fact spherical harmonics are nothing else but homogeneous harmonic polynomials restricted to the sphere [73]. On the other hand, the choice of using tensors instead of equivalent homogeneous polynomials makes some operations elementary and intuitive. This is especially true for our case where only even order symmetric tensors are involved.

In this case tensor products appear formally as quadratic forms on second order tensors (\(\cong\) square matrices) [99], a fact that we will use later in section 5.5.

### 5.3 Finsler Norm on HARDI Higher Order Tensor Fields

We want to show that higher order spherical tensors, such as those fitted to HARDI data, do define a Finsler norm, which can be used in the geometric analysis of this data. We take here as a point of departure a given orientation distribution function (ODF) on a single shell, which if normalized, is a probability density function on the sphere and which can be computed from the data by using one of the methods described in the literature [92][49][102][52][24]. Later we give a precise definition and show explicitly a novel efficient way to compute this ODF in tensor formulation.

For now we take it to be a smooth spherical function which models the probability that a given direction corresponds to a direction of a fiber. We use the heuristics that a high probability of finding a fiber in direction \(y\) corresponds to a larger diffusivity and at the same time to a shorter travel time from the diffusing particle point of view.

As before \(y\) denotes a unit vector while \(y = |y|\mathbf{y}\) is a general vector in \(\mathbb{R}^3\). We denote the \(n\)th (even) order spherical tensor approximating the Möbius inverse of ODF as \(D_n(x, y)\). Denoting ODF as \(\tilde{D}(x, y)\) here, this means that

\[
D_n(x, y) \approx \frac{\overline{D(x, y)}}{|D(x, y)|} \tilde{D}(x, y),
\]

\(5.8\)

where \(\overline{D(x, y)}\) is the average of \(\tilde{D}(x, y)\). We have chosen Möbius inverse instead of multiplicative inverse, because it preserves the average value of the function. In this way we do not introduce different scalings when
discussing diffusion and metric. As a Finsler norm $F(x, y)$, we propose the following

$$F(x, y) = \left( D_n(x, y) \right)^{1/n} = \left( D_{i_1 \ldots i_n}(x) y^{i_1} \ldots y^{i_n} \right)^{1/n}.$$  \hspace{1cm} (5.9)

In the following, we verify that the necessary criteria for a Finsler norm in Definition 4.1 are fulfilled.

1. Differentiability: Since the tensor field $D_n(x)$ is always positive and equivalent to a polynomial of fixed degree, using e.g. polynomial or spline interpolation between the coefficients, we can assume the $D_n(x)$ to be at least twice differentiable w.r.t. $x$. The differentiability of $F$ in $y$ is obvious from (5.9).

2. Homogeneity: Indeed for any $\alpha \in \mathbb{R}_+$, $x \in M$, $v \in T_x M$:

$$F(x, \alpha v) = \left( D_{i_1 \ldots i_n}(x) \alpha y^{i_1} \ldots \alpha y^{i_n} \right)^{1/n} = \alpha F(x, v). \hspace{1cm} (5.10)$$

3. Strong convexity for a tensor based norm: For clarity we introduce a notation for the position $P(i)$ of an index $i$ in a sequence. We compute explicitly the metric tensor from norm (5.9):

$$g_{kl} = \frac{1}{2} \frac{\partial^2 F^2}{\partial y^k \partial y^l}$$

$$= (n - 1) D^{(2-n)/n} D_{i_1 \ldots i_n} y^{i_1} \ldots y^{P(k) - 1} y^{P(k) + 1}$$

$$\ldots y^{P(l) - 1} y^{P(l) + 1} \ldots y^{i_n}$$

$$- (n - 2) D^{(2-2n)/n} D_{i_1 \ldots i_n} y^{i_1} \ldots y^{P(l) - 1} y^{P(l) + 1} \ldots y^{i_n}$$

$$D_{i_1 \ldots k \ldots i_n} y^{i_1} \ldots y^{P(k) - 1} y^{P(k) + 1} \ldots y^{i_n},$$

where we have abbreviated the scalar sum as

$$D := D_{i_1 \ldots i_n} y^{i_1} \ldots y^{i_n}. \hspace{1cm} (5.11)$$

To simplify the expression, we multiply it with a positive scalar $c = D^{(2n-2)/n}$, since it does not affect the positive definiteness. We obtain

$$c \cdot g_{kl} = (n - 1) D D_{i_1 \ldots k \ldots i_n} y^{i_1} \ldots y^{P(k) - 1} y^{P(k) + 1}$$

$$\ldots y^{P(l) - 1} y^{P(l) + 1} \ldots y^{i_n}$$

$$- (n - 2) D_{i_1 \ldots i_n} y^{i_1} \ldots y^{P(l) - 1} y^{P(l) + 1} \ldots y^{i_n}$$

$$D_{i_1 \ldots k \ldots i_n} y^{i_1} \ldots y^{P(k) - 1} y^{P(k) + 1} \ldots y^{i_n}. \hspace{1cm} (5.12)$$
5.3. Finsler Norm on HARDI Higher Order Tensor Fields

We define a special product

\[
\langle u, v \rangle_D = D_{i_1 \ldots i_n} y^{i_1} \ldots y^{i_{n-2}} u^{i_{n-1}} v^{i_n} .
\] (5.14)

Then from (5.13) we have for an arbitrary vector \( v \):

\[
c \cdot g^{ij} v^i v^j = \langle y, y \rangle_D \langle v, v \rangle_D + (n - 2) \left( \langle y, y \rangle_D \langle v, v \rangle_D - \langle y, v \rangle_D \langle y, v \rangle_D \right) .
\] (5.15)

In case (5.14) does define an inner product, the Cauchy-Schwarz inequality says that the last term must be non-negative. Then we have indeed that

\[
g^{ij} v^i v^j > 0 .
\] (5.16)

Thus in case the norm function is a power of an even order tensor, the strong convexity is satisfied if the \( \langle \cdot, \cdot \rangle_D \) is a real inner product i.e. if

\[
\langle v, v \rangle_D = D_{i_1 \ldots i_n} y_1^{i_1} \ldots y_{n-2}^{i_{n-2}} v^{i_{n-1}} v^{i_n} \geq 0 .
\] (5.17)

This is in turn equivalent to requiring that \( D_{i_1 \ldots i_n} y_1^{i_1} \ldots y^{i_{n-2}} \) is a positive definite tensor for every \( y \).

Although we have simplified the strong convexity criterion for tensor based norms, it is not yet a very practical one. However in case of real HARDI data, the zeroth order component is typically dominant and taking a root further reduces radial variation. At least with all HARDI data we have experimented with, we have obtained a strongly convex norm function. Had it not been the case the metric tensors would have become singular.

The main goal of this section was to define Finsler metric tensors \( g_{ij}(x, y) \) corresponding to a given tensorial ODF-field. Generally in Finsler geometry, instead of one metric tensor per spatial point we obtain a bundle of metric tensors depending on parameter \( y \). For illustration of this, see Fig.5.1.
Figure 5.1: Left: A fourth order spherical harmonic (or tensor), fitted to data points, representing the (not homogenized) norm function. Right: Two ellipsoids (green and blue) illustrating the position and direction dependent metric tensors corresponding to the two vectors with corresponding colors.

5.4 Efficient Computing of Single Tensor ODF

In Q-ball imaging the diffusion profile $\Psi$, is described as

$$\Psi(y) = \int_0^\infty P(r|y) dr ,$$  \hspace{1cm} (5.18)

where $P(r|y)$ is the ensemble-average probability that a particle is displaced from initial point $x_0$ to $x_0 + ry$. In [91] it is shown that assuming short diffusion pulses in the scanning protocol, the relation between signal $S(q)$ and $P(r)$ is

$$P(r) = \int_{\mathbb{R}^3} S(q) e^{i2\pi q \cdot r} dq ,$$  \hspace{1cm} (5.19)

where $q$ is the wave vector i.e. a unit vector encoding the direction and pulse duration. Using this relation it is further shown that $\Psi$ in (5.18) can be estimated to be a zeroth order Hankel transformation of itself, which in turn equals the Funk-Radon transform of the Fourier transform of $P(r)$. In [24] this result is applied to the case that the signal is modeled with spherical harmonics and it is shown that the ODF on a single shell can be
approximated as
\[ \Psi(y)|_{r=1} \approx 2\pi \sum_{\ell=0}^{N} \sum_{m=-\ell}^{\ell} c_{\ell m} Y_{m}^{\ell}(y) P_{\ell}(0) , \] (5.20)

where \( P_{\ell} \) denotes the Legendre polynomial of degree \( \ell \).

From now on we denote
\[ \Psi := \Psi(y)|_{r=1} . \] (5.21)

The main benefits of this approach are that one can compute the ODF \( \Psi \) in a fast and robust way using spherical harmonics (SH) and also apply first order approximation of Laplace-Beltrami regularization, which is a generalization of Gaussian smoothing for \( L^2 \) functions on \( \mathbb{R}^2 \) to \( L^2 \) functions on the sphere \( S^2 \) [16][29]. The regularization part we discuss in Chapter 6.

We use this result, but replace the spherical harmonics with a monomial tensor which is more practical for applications using Finsler geometry.

For comparison we recall the procedure of fitting spherical harmonics to HARDI data according to [24]. All functions used below depend on position \( x \) and direction \( y \) (or \( \theta, \varphi \) when confined to the sphere) but for brevity the dependence on \( x \) is suppressed in the following formulae.

This is done as follows.

1. A set of real bases \( \{Y_{i}\}_{i=1}^{N} \) of spherical harmonics up to the desired order is fitted to data sampled on the sphere \( \{S(\theta_k, \varphi_k) \mid k = 1, \ldots, m\} \) using the method of least squares
\[ S(\theta, \varphi) = \sum_{i=1}^{N} c_i Y_i(\theta, \varphi) . \] (5.22)

Here the index \( i \) collectively denotes the indexed pair \((\ell, m)\) as used in (5.20).

2. Each spherical harmonic \( Y_i \) of order \( \ell_i \) is multiplied by a factor \( 2\pi P_{\ell_i}(0) \), where
\[ P_{\ell}(0) = \begin{cases} 0 & \ell \text{ odd} \\ (-1)^{\ell/2} \frac{1 \cdot 3 \cdot 5 \cdots (\ell-1)}{2 \cdot 4 \cdot 6 \cdots \ell} & \ell \text{ even} \end{cases} . \] (5.23)
3. Then
\[
\Psi(\theta, \varphi) := \sum_{i=1}^{N} 2\pi P_{\ell_i}(0)c_i Y_i(\theta, \varphi) .
\] (5.24)

With our tensor approach, we do in principle exactly the same, but in a way that does not require the knowledge or storage of spherical harmonics and allows one to construct Finsler norms related to the local diffusivity.

In our approach the previous steps are replaced by the following:

1. We fit a \(n\)th order tensor \(D\) to the sampled data on the sphere:
\[
S(y) = D(y) = D_{i_1 \cdots i_n} y^{i_1} \cdots y^{i_n} .
\] (5.25)

2. From a tensor \(D(y)\), we compute the harmonic components \(H_n, H_{n-2}, \ldots, H_0\), for which
\[
\triangle^{k/2} H_k(y) = 0, \quad k = n, n-2, \ldots, 0 ,
\] (5.26)
where \(\triangle\) is the Laplace operator and decompose \(D(y)\) as
\[
D(y) = \sum_{k=0}^{n} H_k(y)((y^1)^2 + (y^2)^2 + (y^3)^2)^{(n-k)} .
\] (5.27)

These harmonic components \(H_k(y)\) are in fact eigenfunctions of the Laplace-Beltrami operator on the sphere, with eigenvalues \(k(k+1)\) [73]. This allows us to use the results for Q-ball imaging [24] [92].

The harmonic components can be easily computed using the so-called Clebsch-projection [73], which we discuss in the next section 5.5.

3. Similarly
\[
\Psi(y) := \sum_{i=1}^{N} 2\pi P_{\ell_i}(0)H_i(y) .
\] (5.28)

Thus even the knowledge of a basis of real spherical harmonics is not needed. We remark that an \(n\)th order symmetric tensor in dimension three has
\[
N_o = \binom{n+2}{n} = \frac{(n+2)(n+1)}{2}
\] (5.29)
5.5 Polynomial vs. Monomial Representation of Tensor

A spherical tensor of (even) order $n$ can be represented in two ways. As a single tensor itself or decomposed to harmonic components e.g. via Clebsch-projection, when it becomes a sum of (even) order tensors up to order $n$. Both presentations have their merits.

- A single tensor can be easily homogenized and, with conditions given in section 5.3, modeled as a Finsler-norm. A single tensor best approximating a data set is fast to compute.

- A harmonic decomposition of tensor multiplied with Legendre-coefficients gives directly a good approximation of the ODF $\Psi$. A harmonic decomposition of tensor is easy to smooth with the Laplace-Beltrami operator [73].

For a Finsler approach and efficiency, we would rather work with a single tensor representation

$$D(y) = D_{i_1\ldots i_n} y^{i_1} \ldots y^{i_n}, \quad (5.30)$$

than with the equivalent polynomial expression

$$D(y) = \sum_{k=0}^{n} \tilde{D}_{i_1\ldots i_k} y^{i_1} \ldots y^{i_k}, \quad (5.31)$$

but still exploit the convenient properties of the latter. For an illustration of the hierarchical structure of a fourth order tensor see Fig. 5.2.
Indeed, we can transform a polynomial tensor representation to a monomial one and vice versa, using the fact that our polynomials are restricted to the sphere (5.7).

**From Polynomial to Monomial Expression**

We can take any even order \( n \) tensor and expand it to a \( n + 2 \) order tensor and symmetrize it. Symmetrization of a tensor amounts to a projection to the subspace of symmetric tensors \([99]\). For an even order \((n)\) symmetric tensor \( T \) the multilinear mapping

\[
T : \mathbb{S}^2 \times \cdots \times \mathbb{S}^2 \rightarrow \mathbb{R}, \quad T(y) = T_{i_1 \ldots i_n} y^{i_1} \cdots y^{i_n}, \quad (5.32)
\]

can be equivalently expressed as

\[
\tilde{T}_{\alpha\beta} \tilde{y}^\alpha \tilde{y}^\beta, \quad (5.33)
\]

where \( \alpha, \beta = 1, \ldots, 3^{n/2} \). Here \( \tilde{T} \) is a square matrix with components

\[
\tilde{T}_{\alpha\beta} = T_{\sigma_\alpha(i_1) \cdots \sigma_\alpha(i_{n/2}) \sigma_\beta(i_1) \cdots \sigma_\beta(i_{n/2})}, \quad (5.34)
\]

and

\[
\tilde{y}^\alpha = y^{\sigma_\alpha(i_1)} \cdots y^{\sigma_\alpha(i_{n/2})}, \quad (5.35)
\]

where \( \sigma_1, \ldots, \sigma_{3^{n/2}} \) are the permutations of components of \( y \) corresponding to each term of the outer product \( y \otimes_1 \cdots \otimes_{n/2} y \).
In this matrix-form, for example the transformation of a second order tensor with components $T_{ij}$ to a fourth order tensor with components $T_{ijkl}$ can be illustrated in a very simple way in Fig. 5.3.

The formula for obtaining a fourth order symmetric tensor equivalent to a
given second order Cartesian tensor is thus

\[ T_{ijkl} = \frac{1}{4!} \sum_{\sigma \in S_4} T_{\sigma(i)\sigma(j)} H_{\sigma(k)\sigma(l)}, \]

(5.36)

where \( H \) is the Euclidean metric tensor which in Cartesian coordinates reduces to the identity matrix and \( S_4 \) the symmetric group of all permutations of a set of four elements.

This can be iteratively extended to arbitrary high orders.

**From Monomial to Polynomial Expression**

We can also always transform the monomial expression to a sum of residual monomials that do not contain the factor \( |y|^2 \), using iteratively the Clebsch projection [73].

For example, if we have a monomial tensor \( T_{2n}(y) \) of order \( 2n \), then the corresponding residual part \( Y_{2n}(y) \) is

\[ Y_{2n}(y) = T_{2n}(y) - (\mathbb{D}_{2n}(q) T_{2n})(y), \]

(5.37)

where \( q \) is the dimension of the space and the differential operator \( \mathbb{D} \) is as follows

\[ \mathbb{D}_{2n}(q) = \sum_{k=1}^{n} \frac{(-1)^{k-1} \Gamma(2n - k + \frac{q-2}{2})}{4^k k! \Gamma(2n + \frac{q-2}{2})} |y|^{2k} \Delta^k. \]

(5.38)

Here we show an example of using the formula of Clebsch to project a polynomial of degree \( n \) to a harmonic and a non-harmonic part. Applying the projection iteratively one obtains a decomposition of the original polynomial to a sum of harmonic polynomials of degrees up to \( n \) [73]. Let \( H_n(y) \) be a homogeneous polynomial of degree \( n \). In our case this is

\[ H_n(y) = D_{i_1 \ldots i_k} y^{i_1} \ldots y^{i_k}, \]

(5.39)

where

\[ y = (\sin \theta \cos \varphi, \sin \theta \sin \varphi, \cos \theta), \]

(5.40)

i.e. a homogeneous polynomial restricted to the sphere in \( \mathbb{R}^3 \). The Clebsch projection operator in dimension three is as follows

\[ \mathbb{P}_n := \sum_{k=1}^{[\frac{n}{2}]} \frac{(-1)^{n-1} \Gamma(n - k + \frac{1}{2})}{4^k k! \Gamma(n + \frac{1}{2})} |y|^{2k} \Delta^k, \]

(5.41)
where $\Gamma$ is the Gamma function

$$
\Gamma(z) = \int_0^\infty t^{z-1}e^{-t}dt .
$$

Computing

$$
Y_n(y) = H_n(y) - \mathbb{P}_n(H_n)(y) ,
$$
gives $Y_n(y)$ which is a homogeneous and harmonic polynomial of degree $n$. For proof see [73]. Unless $n = 2$, this procedure is repeated until $H_n$ is totally decomposed into a sum of harmonic homogeneous polynomials.

Next we show an explicit example of computing the $\Psi$ by simply multiplying the $n$th order tensor components fitted to the signal by a matrix. This matrix is obtained by computing once the Clebsch-projection of a dummy-tensor of degree of choice.

For simplicity we show this in degree two, but it extends to all even degrees. Let

$$
S = (S_{11}, S_{12}, S_{13}, S_{22}, S_{23}, S_{33})
$$

be the six components of the symmetric tensor fitted to the HARDI signal e.g. by least squares method. Then the local second order polynomial describing the diffusivity is

$$
H_2(y) = S_{11}y^1y^1 + 2S_{12}y^1y^2 + S_{13}y^1y^3 \\
+ S_{22}y^2y^2 + 2S_{23}y^2y^3 + S_{33}y^3y^3 ,
$$

where $y$ is as in (5.40). Clebsch projection operator (5.41) for a polynomial of degree two (5.45) is:

$$
\mathbb{P}(H_2(y)) = \frac{1}{6}|y|^2\Delta(H_2(y))
= \frac{1}{3}\left((y^1)^2 + (y^2)^2 + (y^3)^2\right) (S_{11} + S_{22} + S_{33})
= \frac{1}{3}(S_{11} + S_{22} + S_{33})
= Y_0(y)
$$
giving the zeroth order harmonic tensor. The harmonic part is

\[ Y_2(y) = H_2(y) - P(H_2(y)) \]

\[ = \left( \frac{2}{3} S_{11} - \frac{1}{3} S_{22} - \frac{1}{3} S_{33} \right) (y^1)^2 \]

\[ + \left( \frac{2}{3} S_{22} - \frac{1}{3} S_{11} - \frac{1}{3} S_{33} \right) (y^2)^2 \]

\[ + \left( \frac{2}{3} S_{33} - \frac{1}{3} S_{22} - \frac{1}{3} S_{11} \right) (y^3)^2 \]

\[ + 2S_{12} y^1 y^2 + 2S_{13} y^1 y^3 + 2S_{23} y^2 y^3. \] (5.47)

Multiplying these harmonic components \( Y_0, Y_2 \) with terms (5.23)

\[ 2\pi P_0(0) = 2\pi \quad \text{and} \quad 2\pi P_2(0) = -\pi, \] (5.48)

we obtain the ODF:

\[ \Psi = 2\pi \left( Y_0(y) - \frac{1}{2} Y_2(y) \right). \] (5.49)

The zeroth order part of \( \Psi \) is

\[ \Psi_0 = \frac{2\pi}{3} (S_{11} + S_{22} + S_{33}), \] (5.50)

which can be expanded to second order tensor by taking the Kronecker product of this scalar and identity matrix \( I_3 \) (similarly for higher order tensors [6]). Thus in matrix form (5.50) is equivalent to

\[ \frac{2\pi}{3} \left( \begin{array}{ccc} S_{11} + S_{22} + S_{33} & 0 & 0 \\ 0 & S_{11} + S_{22} + S_{33} & 0 \\ 0 & 0 & S_{11} + S_{22} + S_{33} \end{array} \right). \] (5.51)

The second order part of \( \Psi \) is:

\[ \Psi_2 = -\pi \left( \frac{2}{3} S_{11} - \frac{1}{3} S_{22} - \frac{1}{3} S_{33} \right) (y^1)^2 \]

\[ + \left( \frac{2}{3} S_{22} - \frac{1}{3} S_{11} - \frac{1}{3} S_{33} \right) (y^2)^2 \]

\[ + \left( \frac{2}{3} S_{33} - \frac{1}{3} S_{22} - \frac{1}{3} S_{11} \right) (y^3)^2 \]

\[ + 2S_{12} y^1 y^2 + 2S_{13} y^1 y^3 + 2S_{23} y^2 y^3. \] (5.52)
In matrix form this

\[-\pi \begin{pmatrix} \frac{2}{3}S_{11} - \frac{1}{3}S_{22} - \frac{1}{3}S_{33} & S_{12} \\ \frac{2}{3}S_{22} - \frac{1}{3}S_{11} - \frac{1}{3}S_{33} & \frac{2}{3}S_{33} - \frac{1}{3}S_{11} - \frac{1}{3}S_{22} \end{pmatrix} \]

(5.53)

Given a symmetric second order tensor \(S(y)\) (5.44) representing the raw HARDI signal, we obtain the symmetric ODF matrix \(\Psi = \Psi_0 + \Psi_2\) as follows

\[
\begin{pmatrix} \Psi_{11} \\ \Psi_{12} \\ \Psi_{13} \\ \Psi_{22} \\ \Psi_{23} \\ \Psi_{33} \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & \pi & 0 & \pi \\ 0 & -\pi & 0 & 0 & 0 & 0 \\ 0 & 0 & -\pi & 0 & 0 & 0 \\ \pi & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -\pi & 0 \\ \pi & 0 & 0 & \pi & 0 & 0 \end{pmatrix} \begin{pmatrix} S_{11} \\ S_{12} \\ S_{13} \\ S_{22} \\ S_{23} \\ S_{33} \end{pmatrix}.
\]

(5.54)

The procedure of obtaining the matrix that transforms an \(n\)th order tensor fitted to signal to a tensor representing ODF is a direct extension of this. These matrices corresponding to higher order tensors have to be computed only once per order. Thus to model a HARDI ODF with a polynomial of order \(n\) one needs only to fit an \(n\)th order tensor (homogeneous polynomial) to the raw signal and multiply it with a pre-computed matrix. This is an efficient alternative for iterative fitting.

### 5.6 Quality Measures for HARDI Fibers

#### Interpretation of Finsler Metric in Applications

We have two heuristic interpretations for the Finsler metric tensor in our applications. When applying Finsler geometry, we deal with norms and inner products that define distances. Then it is intuitive to replace the ODF \(\Psi\) with the Möbius (or multiplicative) inverse of \(\Psi\) in all computations, for larger diffusivity implies shorter travel time from the diffusing particle point of view. Since water molecules in human tissue have constant average velocity, the direct proportionality of (tissue structure induced) distance and diffusion time can be deduced trivially from \(x = vt\).
On the other hand, when we want to work directly with the probabilistic diffusion profile and generalize the streamline tracking to the HARDI case, we adopt the convention $\tilde{F} = \Psi$ and compute the direction dependent diffusion tensor as in (5.11). This diffusion tensor describes a quadratic form corresponding to small perturbations of $y$, which can be seen also from the following second order Taylor expansion of $F^2$ with fixed $x$. For brevity we denote $F' = \frac{d}{dy} F$.

\[
F^2(y + \delta) = F^2(y) + (F^2(y))' \delta + \frac{1}{2} \left( F^2(y) \right)'' \delta^2 + O^3
\]

\[\approx F^2(y) + 2F(y)F'(y)\delta + \frac{1}{2} \left( F^2(y) \right)'' \delta^2,
\]

from which we get

\[
\frac{1}{2} \left( F^2(y) \right)'' \approx \frac{(F(y + \delta) - F(\delta))^2}{\delta^2}.
\]

Connectivity

We discuss two quality measures for curves in HARDI that can be used to any curve regardless of the tracking method. We motivate our approach on an analytical example, where the optimal paths are geodesics and begin by briefly recalling geodesics in Finsler geometry (Sec. 4.5).

Similar to the Riemannian case, in Finsler setting a geodesics $\gamma(t)$ is a curve that satisfies the Euler-Lagrange equations of the length integral. This gives us a local condition according to which $\gamma$ has vanishing geodesic curvature (4.3) [86]. This amounts to equation

\[
\dddot{\gamma}^i(t) + 2G^i(\gamma(t), \dot{\gamma}(t)) = 0, \quad (5.57)
\]

where

\[
G^i(x, y) = \frac{1}{4} g^{il}(x, y) \left( \frac{\partial F^2}{\partial x^k \partial y^l}(x, y) y^k - \frac{\partial F^2}{\partial x^l}(x, y) \right),
\]

is the so-called geodesic coefficient. Using the relationship of the norm and the bilinear form

\[
g_{ij} = \frac{1}{2} \frac{\partial^2 F^2}{\partial y^i \partial y^j},
\]

(5.59)
we see that
\[
2G^i = \frac{1}{2} g^{il} \left( \frac{\partial}{\partial x^k} \left[ \frac{\partial}{\partial y^l} g_{jn} y^j y^n \right] - \frac{\partial}{\partial x^l} \left[ g_{jk} y^j y^k \right] \right)
\]
\[
= \frac{1}{2} g^{il} \left( 2 \frac{\partial g_{jl}}{\partial x^k} y^j y^k - \frac{\partial g_{jk}}{\partial x^l} y^j y^k \right) \quad (5.60)
\]
\[
= \frac{1}{2} g^{il} \left( \frac{\partial g_{jl}}{\partial x^k} + \frac{\partial g_{kl}}{\partial x^j} - \frac{\partial g_{jk}}{\partial x^l} \right) y^j y^k ,
\]
which reveals that the geodesic equation in the Finslerian case is formally identical to that in the Riemann geometry

\[
\dot{\gamma}^i + \Gamma^i_{jk} y^j y^k = 0, \quad (5.61)
\]
where \( \Gamma^i_{jk} \) is the last term in (5.60). Note that, unlike in the Riemannian case, the symbols \( \Gamma^i_{jk} \) depend on \( y \).

We take an analytic norm field in \( \mathbb{R}^2 \) as an example, but the situation can be directly extended to \( \mathbb{R}^3 \). Let us take as a convex norm function at each spatial position

\[
F(\varphi) = (\cos 4\varphi + 4)^{\frac{1}{4}} = (5 \cos^4 \varphi + 2 \cos^2 \varphi \sin^2 \varphi + 5 \sin^4 \varphi)^{\frac{1}{4}} \quad (5.62)
\]
This is an example of a fourth order spherical tensor. Such a tensor field could represent a dense field of orthogonally crossing fibers (cf. Fig.5.5). From the fact that \( F \) has no \( x \)-dependence we conclude that the geodesic coefficients vanish and that the geodesics coincide with the Euclidean geodesics \( \gamma(t) = (t \cdot \cos \varphi, t \cdot \sin \varphi) \), i.e. straight lines. However the so-called connectivity of a geodesic [7][80] is relatively large, only in cases, where the directional norm function is correspondingly small. In Finsler setting the connectivity measure \( m(\gamma) \) is:

\[
m(\gamma) = \frac{\int \sqrt{\eta_{ij} \dot{\gamma}^i \dot{\gamma}^j dt}}{\int \sqrt{g_{ij}(\gamma, \dot{\gamma}) \dot{\gamma}^i \dot{\gamma}^j dt}}, \quad (5.63)
\]
where the \( \eta_{ij}(\gamma) \) represents the covariant Euclidean metric tensor which in Cartesian coordinates reduces to the identity matrix, \( \dot{\gamma}(t) \) the tangent to the curve \( \gamma(t) \) and \( g_{ij}(\gamma, \dot{\gamma}) \) the Finsler-metric tensor (which depends not only on the position on the curve but also on the tangent of the curve).
For illustration we compute explicitly the metric tensors $g$, using Cartesian coordinates:

$$g = \frac{1}{2(4 + \cos 4\varphi)^{3/2}} \begin{pmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{pmatrix},$$ (5.64)

where

\begin{align*}
g_{11} &= 5(6 + 3 \cos 2\varphi + \cos 6\varphi), \\
g_{12} &= g_{21} = -12 \sin 2\varphi^3, \\
g_{22} &= -5(-6 + 3 \cos 2\varphi + \cos 6\varphi).
\end{align*}

We denote here the indicatrix (4.7) as the set $u$, i.e. the set for which $F(y) = 1 \forall y \in u$. Here it is the parametric set

$$u = \frac{1}{(\cos 4\varphi + 4)^{1/4}} (\cos \varphi, \sin \varphi), \quad \varphi \in [0, 2\pi]$$ (5.65)

The strong convexity criterion, which suffices to be satisfied on the indicatrix [10], is as follows

$$\ddot{u}^1u^2 - \dot{u}^1\dot{u}^2 > 0,$$ (5.66)

where $\dot{u} := \frac{d}{d\varphi}u$ etc. For illustration of this criterion see Fig. 5.4. This criterion is satisfied since

$$\frac{\ddot{u}^1\dot{u}^2 - \dot{u}^1\ddot{u}^2}{\dot{u}^1u^2 - u^1\dot{u}^2} = \frac{13 + 8 \cos 4\varphi}{(4 + \cos 4\varphi)^2} > 0.$$ (5.67)

The connectivity measure (5.63) for a (Euclidean) geodesic $\gamma$ can be computed analytically:

$$m(\gamma) = \frac{\int dt}{\int (4 + \cos(4\varphi))^{1/4} dt},$$ (5.68)

which gives the maximal connectivities in directions $\{\pi/4, 3\pi/4, 5\pi/4, 7\pi/4\}$, as expected. See Fig. 5.5 for an illustration. We observe that on such a norm field the Riemannian (DTI) framework would result in Euclidean geodesics and constant connectivity over all geodesics thus revealing no information at all of the angular heterogeneity.

### Flag Curvature

When in possession of a metric tensor field (such as DTI data), one can compute the geodesic deviation that measures whether the neighboring
Figure 5.4: The dashed line in pink is the norm function and the blue line is the indicatrix. We have computed derivatives at the yellow spots. Red arrow is $u$, green arrow is $\dot{u}$ and purple arrow is $\ddot{u}$. The strong convexity requires that the dipods $(\dot{u}, u)$ and $(\dot{u}, \ddot{u})$ have different orientations as in this picture.

geodesics, whose initial tangents are confined in a plane, deviate from each other or merge together. This can be computed directly from the tensor data without computing geodesics first. (For illustration of geodesic deviation recall the left most picture of Fig. 3.6.) We show here how to compute geodesic deviation as a function of position and direction in the HARDI case.

As a generalization of the Riemannian sectional curvature (which is a generalization of the famous Gauss curvature to higher dimensions), in Finsler geometry we have the so-called flag curvature. In notation of [21][86], the flag curvature $K(y, V)$ is defined as

$$K(y, u) := \frac{\langle R_y(u), u \rangle_{g(y)}}{\langle y, y \rangle_{g(y)} \langle y, y \rangle_{g(y)} - \langle y, u \rangle_{g(y)} \langle y, u \rangle_{g(y)}}, \quad (5.69)$$
Figure 5.5: Left: A field of fourth order spherical harmonics \( \cos 4\varphi + 4 \) representing the norm. In the middle of the figure, the ODF profiles (a Moebius-inverse of the norm) are indicated and some best connected geodesics are drawn. Right: 50 ellipses representing metric tensors corresponding to directions \( \varphi = \frac{2\pi}{50}i \), \( (i = 1, \ldots, 50) \) of the norm function \( F(\varphi) \), and an ellipse with thick boundary corresponding to the metric tensor in direction \( \varphi = \frac{\pi}{4} \). The thick red curve is the homogenized norm function \( F(\varphi) ^{\frac{1}{2}} \).

where \( g(y) \) is the generalized metric tensor defined by position \( x \) and direction \( y \). Brackets \( \langle \, , \rangle_{g(y)} \) denote the inner product w.r.t. metric tensor \( g_y \) and \( R_y(u) \) is the Riemann tensor determined by vectors \( y, u \in T_xM \), which we define shortly. Although these do depend on \( x \) we suppress \( x \) from notation to avoid clutter.

The Riemann tensor \( R_y \) maps the vector \( u^s \frac{\partial}{\partial x^s} \) to a vector:

\[
R_y(u) = R^i_k(x,y)u_k \frac{\partial}{\partial x^i} ,
\]

where

\[
R^i_k = 2 \frac{\partial G^i}{\partial x^k} - y^j \frac{\partial^2 G^i}{\partial x^j \partial y^k} + 2G^j \frac{\partial^2 G^i}{\partial y^j \partial y^k} - \frac{\partial G^i}{\partial y^j} \frac{\partial G^j}{\partial y^k} ,
\]

with coefficients \( G^i \) defined as:

\[
G^i = \frac{1}{4}g^{il} \left( \frac{\partial^2 F^2}{\partial x^k \partial y^l} y^k - \frac{\partial F^2}{\partial x^l} \right). 
\]

Flag curvature \( K(y,u) \) is a measure for geodesic deviation within the plane spanned by vectors \( y \) and \( u \) [10]. Neural fibers are generally not expected
5.7. Streamline Tracking of HARDI Fibers

To be planar, thus we want to look at all geodesics that are slight perturbations of the fiducial geodesic, rather than only those restricted to one plane. In analogy to the Riemannian case, we do this by considering the average of the geodesic deviations w.r.t. all planes containing the initial vector of the fiducial geodesic. In the Riemannian case, this average is equivalent to the Ricci or mean curvature. Ricci curvature in Finsler geometry is defined similarly in dimension $n$. Using Eq. 6.10, this is [21]

$$Ric(y) = R(y)^m_m.$$  

(5.73)

If $Ric(y)$ is positive, in average, the geodesics will locally merge together. If $Ric(y)$ is zero, in average, the geodesics will diverge linearly proportionally to the angle of perturbation. If $Ric(y)$ is negative, in average, the geodesics will diverge [20].

5.7 Streamline Tracking of HARDI Fibers

The most popular way to track axon(bundle)-like curves in DTI data is to follow the principal eigenvectors of diffusion tensors $D_{ij}(x)$. This approach requires that the curve $c : [0, 1] \rightarrow \mathbb{R}^3$ satisfies the following equations

$$
\begin{align*}
\dot{c}(t) &= \arg \max_{|h|=1} \{ D_{ij}(c(t)) h^i h^j \} , \\
c(0) &= p , \\
\dot{c}(0) &= \arg \max_{|h|=1} \{ D_{ij}(c(0)) h^i h^j \} ,
\end{align*}
$$

(5.74)

The solution to this equation is a geodesic from a point $p$ to a sphere $S^2(p, \varepsilon)$ for some small $\varepsilon$, but not necessarily a geodesic between $c(0)$ and $c(1)$. This tracking scheme can be extended to the HARDI framework using Finsler geometry as follows

$$
\begin{align*}
\dot{c}(t) &= \arg \max_{|h|=1} \{ D_{ij}(c(t), \dot{c}(t)) h^i h^j \} , \\
c(0) &= p ,
\end{align*}
$$

(5.75)

where

$$
\dot{c}(t) = \lim_{\delta \to 0} \frac{c(t) - c(t - \delta)}{\delta} .
$$

(5.76)
and the second order tensor \( D_{ij}(c(t), \dot{c}(t)) \) is computed from an \( n \)th order tensor \( T_n(c(t), \dot{c}(t)) \) (approximating ODF) as follows

\[
D_{ij}(c(t), \dot{c}(t)) = \frac{1}{2} \frac{\partial^2 ((T_n(c(t), y))^{1/n})^2}{\partial y^i \partial y^j} |y = \dot{c}(t) .
\] (5.77)

The practical implementation is very similar to that of the streamline tracking in DTI setting. A first difference is that since the diffusion profile is more complex than an ellipsoid, there is no unique initial principal eigenvector at the initial point of tracking. Instead we start tracking to all gradient directions. The user can define gradients to be e.g. the points of tessellations of order of choice. Here we used the second order tessellation (zeroth order being the icosahedron) with 54 gradients. A second difference is that there is no unique second order diffusion tensor per point. Instead there is a unique second order local diffusion tensor corresponding to the direction of arrival (5.76) at the point. This local diffusion tensor is computed as in (5.11). Because our ODF is a \( 1/n \)th power of a single tensor, the formula for a local diffusion tensor is easily computed beforehand. For example if we have a fourth order tensor

\[
D_4(y) := D_{ijkl} y^i y^j y^k y^l
\] (5.78)

describing the diffusion profile, the Finsler metric tensor corresponding to the homogenized form

\[
F(y) := (D_4(y))^{1/4}
\] (5.79)

is

\[
g_{ij}(y) = \frac{1}{2} \frac{\partial^2 F^2}{\partial y^i \partial y^j}
= 3(D_4)^{-1/2} D_{ijqp} y^q y^p
- 2(D_4)^{-3/2} D_{ijkl} y^i y^j y^k y^l D_{ijmn} y^m y^n
\] (5.80)

Then we have a bi-linear form at every position such that filling in the vector \( y \) of direction of arrival, it returns a local diffusion tensor.

If the direction of arrival is close to the principal eigenvector of this local diffusion tensor, we take a step forward, if not we stop.

The familiar anisotropy criterion for DTI, which stops tracking if the diffusion tensor is isotropic, can be applied to this local diffusion tensor as
5.7. Streamline Tracking of HARDI Fibers

well. Thus in a nutshell the algorithm is simply as follows: If the tangent of the streamline is in alignment with the principal eigenvector defined by the local diffusion tensor (corresponding to the tangent) and the fractional anisotropy of the local diffusion tensor is high enough, proceed, otherwise stop.

5.7.1 Experimental Results

Simulated Data

We computed Finsler streamlines in tensor fields that simulate two fiber bundles crossing at angles of 30 and 65 degrees. The tensors in the crossing area are generated using the Gaussian mixture model [93] assuming signal with b-value 1000. Noisy images were generated by adding Rician noise with SNR around 15. We solved the Finsler streamlines (5.74) using a third order Runge-Kutta model for integration [40]. If the fractional anisotropy (FA) became less than 0.2 or if the inner product of the tangent of the curve and the local principal eigenvector became less than 0.1, the tracking stops. These parameters are similar to those used in the standard DTI-fiber tracking. We have colored the fibers using a temperature map according to their connectivity (5.63) so that fibers with highest (lowest) connectivity are colored red (blue). See Figs. 5.7 and 5.6.

Real Data

We used a HARDI scan of human brain with b-value 1000, and 132 gradient directions. As initial points we used 30 voxels in the area where the major fiber bundles called Corpus Callosum and the Corona Radiata are known to cross each other. The parameters used are identical to those used in simulated data except that the critical value of FA was set to 0.09. For comparison we added also the corresponding second order tensor field and a result of standard DTI-tracking in Fig. 5.8. Finsler streamlines in Fig. 5.9 indeed do show crossings in the regions where this is to be expected. In Fig. 5.10 a tracking result on a larger initial point set is shown with a zoom up that shows how the fibers actually cross each other.
Figure 5.6: Left: Streamlines in a noise free tensor field simulating a 65° crossing of fiber bundles. Streamlines are colored according to their connectivity. High (low) connectivity is color coded as red (blue). Right: Same field with Rician noise added.

Figure 5.7: Left: Streamlines in a noise free tensor field simulating a 30° crossing of fiber bundles. Note that smaller angle results in kissing streamlines. Right: Same field with Rician noise added.

Figure 5.8: Second order tensor field on real data and DTI-streamlines with black seed points. In DTI, intersecting fibers are impossible since there is only one tangent vector occupying a point. Apparent crossings are due to the projection of spatial curves to image plane.
Figure 5.9: Fourth order tensor glyphs on real HARDI data and Finsler streamlines with seed points. Although these streamlines are generally not on the same plane, they do actually cross each other in a way that is in accordance with the anatomical knowledge of human brain [55].

Figure 5.10: Left: Finsler-fibers on a human brain data. Right: A zoom in to the crossings of the corpus callosum and the corona radiata (red and blue). The yellow fibers resembling the cingulum run over the red fiber layer (the corpus callosum).
"Speed is for parrots and machines; Human beings work better at a more deliberate pace.” Laurence C. Young
6.1 Diffusion Tensor Interpolation Problem

Introduction

If we want to integrate, differentiate or visualize some continuum in a given discrete space of symmetric positive definite (SPD) matrices, we typically have to do interpolation among these matrices. Taking (weighted) means is one part of the task. Given elements \( \{x_1, x_2, \ldots, x_n\} \) of a metric space, a mean of these elements is another element of the space that can be "connected to every element with minimal overall cost". The so-called Fréchet-mean is a solution \( x_0 \) to the equation

\[
x_0 = \arg\min_x \sum_{i}^{n} d(x, x_i)^2 ,
\]

(6.1)

and depends on the given distance measure \( d( , , ) \).

For a set of positive real numbers \( (s_1, \ldots, s_N) \), we have for example the arithmetic mean

\[
\frac{1}{N} \sum_{i=1}^{N} s_i ,
\]

(6.2)

which gives the solution to (6.1) with respect to the Euclidean metric \( d(s_i, s_j) = \sqrt{(s_i - s_j)^2} \) and the geometric mean

\[
\sqrt[N]{s_1 \cdots s_N} ,
\]

(6.3)

which gives the solution to (6.1) with respect to the hyperbolic metric: \( d_h(s_i, s_j) = |\log(s_i) - \log(s_j)| \).

An algebraic definition for the geometric mean of two elements is that it is a solution \( x \) to the problem

\[
x^2 = ab \quad \text{or equivalently} \quad xa^{-1}x = b .
\]

(6.4)
Some Formulae for Means for SPD-Tensors

In the space of invertible \( n \times n \) matrices (i.e. the general linear group \( GL(n) \)), the Frobenius inner product is defined as

\[
\langle A, B \rangle_F = \text{tr}(A^T B).
\] (6.5)

From this we obtain the so-called Frobenius norm

\[
||A||_F = (\langle A, A \rangle_F)^{1/2}.
\] (6.6)

The distance between matrices \( A \) and \( B \) is then

\[
d_F(A, B) = ||A - B||_F.
\] (6.7)

Solving the Frechet mean for the Frobenius distance, we obtain the componentwise arithmetic mean of matrices

\[
\mu_A = \frac{1}{2} (A + B).
\] (6.8)

The space of symmetric matrices \( Sym(n) \) forms an additive group and a vector space, but the space of symmetric positive definite (DTI-) tensors \( Sym^+(n) \) does not form the latter. Thus if we want to emphasize the importance of positive definiteness we might want to avoid using the previous distance measure in \( Sym^+(n) \), since it will give in some cases a shorter distance to a non-SPD matrix than to some SPD matrix. In addition, the arithmetic mean is not invariant w.r.t. inversion, unlike the means we introduce in the following.

There is a bijective differentiable mapping from \( Sym(n) \) to \( Sym^+(n) \), namely the matrix exponential:

\[
\text{Exp}(A) = \sum_{k=0}^{\infty} \frac{A^k}{k!}.
\] (6.9)

With this mapping, we can consider \( Sym^+(n) \) as a Riemannian manifold with \( Sym(n) \) as the tangent space. The geometric mean of these matrices has been studied by several authors [81][57][13]. In [69] a Riemannian inner product on the tangent space \( Sym(n) \) at \( P \in Sym^+(n) \), is defined as

\[
\langle A, B \rangle_P = \text{tr}(P^{-1}AP^{-1}B),
\] (6.10)
and the respective norm
\[ ||A||_P = \langle A, A \rangle_P^{1/2}. \] (6.11)

The geometric mean \( X \) of matrices \( A \) and \( B \) is a solution to
\[ X A^{-1} X = B \quad \text{or} \quad X B^{-1} X = A. \] (6.12)

When \( A, B \in \text{Sym}^+(n) \), this gives us a unique mean:
\[ \mu_G(A, B) = A^{1/2}(A^{-1/2}BA^{-1/2})^{1/2}A^{1/2}, \] (6.13)
which is equivalent to
\[ \mu_G(A, B) = A(A^{-1}B)^{1/2} = B(B^{-1}A)^{1/2}. \] (6.14)

In [3] the so-called Log-Euclidean mean of matrices \((M_1, \ldots, M_N)\) is introduced
\[ \mu_{LE} = \text{Exp} \left( \frac{1}{N} \sum_{i=1}^{N} \text{Log}(M_i) \right). \] (6.15)

There are of course many other interesting means defined for matrices, for example the Kullback-Leibler mean defined for covariance matrices of Gaussian distributions [70]. In Fig. 6.1 we illustrate some differences between componentwise (6.8), geometric (6.14) and log-Euclidean (6.15) means of pairs of matrices. We show two tensors with colors red and yellow, and their componentwise, geometric and log-Euclidean means in orange color. We notice that in the second row, where we have two planar (pancake-like) diffusion ellipsoids with identical shapes but with different orientations, all mean ellipsoids have linear (cigar-like) shape, with principal eigenvector being the axis of rotation of two planar ellipsoids. The two rows at the bottom of the figure show the most striking difference between the componentwise mean and the other two. The shape of the componentwise mean ellipsoid is affected also by the size of the given tensors. The other mean ellipsoids are less sensitive to the size differences. This may be important for example in a region close to the boundary of ventricles in a DTI image, where diffusion tensors have determinants very different in magnitudes. Another important fact is that the geometric mean is not readily extendable to a mean of more than two tensors, due to the non-commutativity of matrices. In most of the computations done concerning this thesis, for simplicity and speed, we have used the componentwise linear interpolation, but there are also more sophisticated methods e.g. PDE based interpolation methods for SPD tensor fields [97].
Figure 6.1: Columns from left to right: Matrix $A$, $\mu_A(A, B)$, $\mu_G(A, B)$, $\mu_{LE}(A, B)$, matrix $B$. 
6.2 A Scale Space for Diffusion Tensors

Diffusion tensor images contain noise. Therefore it is important to have means to smoothen the images and suppress the effect of noise. A well established way to smoothen scalar images is to apply Gaussian blurring leading to what is called the Gaussian scale space of the image. Scale space theory [45][62] is a framework for multiscale image representation, developed by researchers in computer vision, which is motivated by physics as well as biologic vision. The basic idea is that if a computer is given a task to automatically interpret an image or a scene, it has to have some knowledge of the relevant scales. But since this knowledge is not present in the images themselves, a computer needs to do image processing at multiple scales. Scale space theory is widely studied and has many interesting applications [9][50][79]. A brief review on the history of Gaussian scale space in image analysis can be found in [98]. In this section we study scale space on symmetric positive definite tensor valued images.

6.2.1 Pseudo-Linear Scale Space

The so-called pseudo-linear scale space for SPD tensors was introduced in [30]. Pseudo-linear scale space on SPD-tensors satisfies the requirement of commutativity of two operations: Gaussian derivative and inverse. Such commutativity issue is a natural one in numerical computations of Riemannian quantities that contain derivatives of inverses, such as the Ricci tensor. In the absence of this commutativity we have to decide which operation we should apply first.

We recall that a zeroth order Gaussian derivative is an operation that blurs a \( \mathbb{L}_2 \) function in \( \mathbb{R}^n \) at a given scale \( \sigma \), evolving according to the diffusion equation w.r.t. parameter \( \sigma \).

Let us first consider a positive scalar-valued piecewise continuous function \( f \in \text{Sym}^+(1) \), such that \( y \cdot f(x) \cdot y > 0 \forall x, y \in \mathbb{R} \). We take the inverse \( g = f^{-1} \), so that \( f(x) \cdot g(x) = g(x) \cdot f(x) = 1 \). We smoothen \( f \) and \( g \) by convolving them with a Gaussian at a scale \( \sigma \). The normalized Gaussian is

\[
\phi_\sigma(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left[ -\frac{1}{2} \frac{||x||^2}{\sigma^2} \right] \quad \text{with } \sigma > 0 \quad (6.16)
\]
We denote the convolution in a standard way
\[
(f * g)(x) = \int_{-\infty}^{\infty} f(x - y)g(y)dy .
\] (6.17)

For the blurred functions the following is not always true
\[
(\phi_\sigma * f) \cdot (\phi_\sigma * g) = 1 .
\] (6.18)

This is easy to see e.g. when
\[
f(x) = \exp(x), \ g(x) = \exp(-x) \text{ and } \sigma = 1 ,
\] (6.19)
then
\[
(\phi_\sigma * f)(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp\left(-\frac{(x - \tau)^2}{2}\right) \exp(\tau)d\tau = \exp\left(\frac{1}{2} + x\right) ,
\] (6.20)
and
\[
(\phi_\sigma * g)(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp\left(-\frac{(x - \tau)^2}{2}\right) \exp(-\tau)d\tau = \exp\left(\frac{1}{2} - x\right) .
\] (6.21)

Clearly direct blurring does not preserve the inverse property.

As a remedy for this, we look at representations of $Sym^+_1$ in $\mathbb{R}$
\[
\log : Sym^+_1 \to \mathbb{R} : f(x) \mapsto \log(f(x)).
\] (6.22)

In $\mathbb{R}$ equipped with addition we have $g = f^{-1}$ again, for
\[
\log(f(x)) + \log(g(x)) = \log(f(x) \cdot g(x)) = 0 .
\] (6.23)

Convolving $f$ and $g$ with a Gaussian
\[
u_\sigma(x) = (\phi_\sigma * \ln(f))(x) , \quad u_\sigma(x) = (\phi_\sigma * \ln(g))(x) ,
\] (6.24)
and applying mapping
\[
\exp : Sym(1) \to Sym^+_1 : u_\sigma \mapsto \exp(u_\sigma) ,
\] (6.25)
the blurred logarithms of functions return to the space of positive scalar functions again.
Thus by replacing the Gaussian smoothing

$$\phi_\sigma \ast f, \quad (6.26)$$

with a pseudo-linear Gaussian smoothing

$$\exp(\phi_\sigma \ast \ln(f)), \quad (6.27)$$

the inversion commutes with the blurring

$$\exp(\phi_\sigma \ast \ln(f)) \ast \exp(\phi_\sigma \ast \ln(g)) = 1. \quad (6.28)$$

This result holds also if we replace $\phi_\sigma$ with derivatives of $\phi_\sigma$ due to commutativity of derivative and smoothing operators[42].

Here we extend the pseudo-linear scale space to $Sym^+(3)$ according to [30]. For this we need to use a generalization of derivatives to matrices. As mentioned before a matrix $M$ in $Sym^+(3)$ is of form $M = e^A$, where $A \in GL(n)$ is symmetric. Matrix exponentials can be differentiated using the so-called Frechet derivative[54][65][8]. The Frechet derivative of a matrix exponential of $M \in GL(n)$ in direction of matrix $N$ is defined as

$$D(N, M) = \int_0^1 e^{tM}Ne^{(1-t)M} dt. \quad (6.29)$$

This can be applied for example to study the sensitivity of the matrix exponential to perturbations in its matrix argument [74].

Before we define a scale space differential of matrix exponential, we take a look at two different ways of taking matrix logarithms, since we will need this operation for the derivatives. Let $M = e^N \in Sym^+(n)$ for some unknown $N$. It has a singular value decomposition (SVD)

$$M = XDX^{-1}, \quad (6.30)$$

and

$$\log(M) = \sum_{k=1}^n \log(\lambda_k)V_kV_k^T, \quad (6.31)$$

where $\lambda_k$ are the eigenvalues, and $V_k$ the eigenvectors of $M$. To compute the matrix logarithm via SVD, one has to solve both the eigenvalues and the eigenvectors.
As an alternative to this we can use the so-called Sylvester’s formula \[44\]

\[
\log(M) = \sum_{i=1}^{n} \log(\lambda_i) A_i ,
\]

where

\[
A_i = \prod_{j=1, j \neq i}^{n} \frac{1}{\lambda_i - \lambda_j} (M - \lambda_j I_n) .
\]

(6.33)

We give here an elementary proof that the right hand sides in (6.30) and in (6.32) are indeed the same.

**Lemma 6.1.** Let \( A_i \) be as defined in (6.33) and \( B_i = V_i V_i^T \) linear operators in \( \mathbb{R}^n \).

\[
A_i \equiv B_i .
\]

(6.34)

Proof. Let \( W \) be any vector in \( \mathbb{R}^n \). The unit eigenvectors \( V_i \) form an orthonormal basis and we can express \( W \) as \( W = \sum \alpha_i V_i \). The \( k \)th component of the product of matrix \( B_i \) and vector \( W \) is given by

\[
(B_i \cdot W)^k = \sum_{l=1}^{n} V_i^k V_i^l \cdot W^l = \alpha_i V_i^k .
\]

(6.35)

Thus also

\[
B_i \cdot W = \alpha_i V_i .
\]

(6.36)

For the operator \( A \), we have

\[
A_i \cdot W = \prod_{j=1, j \neq i}^{n} \frac{1}{\lambda_i - \lambda_j} (M - \lambda_j I_n) \sum_{i=1}^{n} \alpha_i V_i
\]

\[
= \frac{1}{\lambda_i - \lambda_1} (M - \lambda_1 I_n) (\alpha_i V_i + \alpha_1 V_1)
\]

\[
= \frac{1}{\lambda_i - \lambda_1} \alpha_i (M V_i - \lambda_1 V_i)
\]

\[
= \alpha_i \frac{1}{\lambda_i - \lambda_1} ((\lambda_i - \lambda_1) V_i)
\]

\[
= \alpha_i V_i ,
\]

since the terms

\[
(M - \lambda_j I_n) \alpha_j V_j = \alpha_j (\lambda_j - \lambda_j) V_j = 0 ,
\]

(6.37)
for all $j \neq i$. Because $W$ and $i$ were arbitrary, we have that
\[
A_i = B_i .
\]
\[
(6.38)
\]

We return to the Gaussian derivatives of SPD matrices. In [30] scale space extensions of the derivative operator up to second order (6.29) are derived. The first order scale space derivative for SPD tensor $e^M$ is
\[
\frac{\partial}{\partial \alpha} e^M(x) = \int_0^1 e^\left[(1-\tau)M(x)\right] \left( \frac{\partial}{\partial \alpha} \phi_\sigma \right) * M \left( x \right) e^{\tau M(x)} d\tau ,
\]
\[
(6.39)
\]
where $\alpha = 1, 2, 3$ denotes the index of Euclidean basis vector. Since Gaussian derivatives have a fast implementation [94], (6.39) can be implemented to perform in a reasonable time.

With the pseudo-linear derivative operator (6.39) we can in principle derive expressions for standard Riemannian quantities like Ricci tensor as follows. Using the fact that for Christoffel symbols in Riemannian setting:
\[
\Gamma^i_{il} = \frac{\partial \log \sqrt{g}}{\partial x^l} ,
\]
\[
(6.40)
\]
where $g$ is the determinant of the metric tensor $G$, it can be shown that Ricci tensor can be expressed as
\[
R_{ij} = \frac{\partial^2 \log \sqrt{g}}{\partial x^i \partial x^j} - \frac{\partial}{\partial x^k} \Gamma^k_{ij} + \Gamma^m_{ik} \Gamma^k_{mj} - \Gamma^m_{ij} \frac{\partial \log \sqrt{g}}{\partial x^m} .
\]
\[
(6.41)
\]
Whereas the traditional form is:
\[
R_{ik} = \Gamma^l_{ik} \Gamma^m_{ml} - \Gamma^l_{mk} \Gamma^m_{il} + \frac{\partial}{\partial x^m} \Gamma^m_{ik} - \frac{\partial}{\partial x^i} \Gamma^m_{mk} .
\]
\[
(6.42)
\]
In both cases, we need the expression for scale space Christoffel symbols:
\[
\Gamma^l_{ik} = \frac{1}{2} g^{ls} \left( \int_0^1 e^{(1-\alpha)(\log G*\phi_\sigma)} \phi_{\sigma,k} * (\log G) e^{\alpha(\log G*\phi_\sigma)} d\alpha \right)_{si} \\
+ \frac{1}{2} g^{ls} \left( \int_0^1 e^{(1-\alpha)(\log G*\phi_\sigma)} \phi_{\sigma,i} * (\log G) e^{\alpha(\log G*\phi_\sigma)} d\alpha \right)_{ks} \\
- \frac{1}{2} g^{ls} \left( \int_0^1 e^{(1-\alpha)(\log G*\phi_\sigma)} \phi_{\sigma,s} * (\log G) e^{\alpha(\log G*\phi_\sigma)} d\alpha \right)_{ik} ,
\]
where expression $\phi_{\sigma,i}$ means a first order derivative w.r.t. $e_i$ of Gaussian function at scale $\sigma$. In formula (6.41) we have to take derivatives of Christoffel symbols only once, which reduces the amount of calculations. We may also use the fact that

$$\det(e^A) = e^{\text{tr}A}. \quad (6.43)$$

Then formula (6.41) becomes:

$$R_{ik} = \frac{\partial^2 \log \sqrt{e^{\text{tr}(\log(G)\ast \phi_\sigma)}}}{\partial x^i \partial x^k} - \frac{\partial}{\partial x^k} \Gamma_{ij}^k + \Gamma_{ik}^m \Gamma_{mj}^k - \Gamma_{mij} \frac{\partial \log \sqrt{e^{\text{tr}(\log(G)\ast \phi_\sigma)}}}{\partial x^m}$$

$$= \frac{\partial^2 \text{tr} (\log(G) \ast \phi_\sigma)}{2\partial x^i \partial x^k} - \frac{\partial}{\partial x^k} \Gamma_{ij}^k + \Gamma_{ik}^m \Gamma_{mj}^k - \Gamma_{mij} \frac{\partial \text{tr} (\log(G) \ast \phi_\sigma)}{2\partial x^m}.$$

Another interesting way to apply diffusion equation to tensor valued images would be to evolve the tensor field simulating the Ricci flow [90], which is a counter-part of the heat equation on metric tensor fields, although this would be computationally quite an intensive task.

### 6.2.2 Experiments

Here we have experimented on how different SPD tensor fields change under blurring, when we use three different methods. The three different methods we use are linear componentwise Gaussian blurring with scale $\sigma$, pseudo-linear blurring with scale $\sigma$ and pseudo-linear blurring with scale $e^\sigma$. The following examples are two dimensional tensor fields. From Fig. 6.2, 6.3 and 6.4 we see that in pseudo-linear blurring the dominant shape does not follow that of the biggest neighboring tensor. This is because the averaging is done in the logarithmic space $\text{Sym}(2)$ where differences in magnitudes are smaller than in $\text{Sym}^+(2)$. Therefore pseudo-linear blurring seems to restore quite well also data with added noise. But if we compare last two columns from the left in Fig. 6.4, the pseudo-linear blurring looses the thin vertical line whereas the linear blurring still shows the line. The three dimensional tensor fields consist of tensors with similar magnitude and the blurred versions are rather similar. For scalar images a pseudo-linear scale space was introduced in [34] as one between the linear and the morphological scale spaces[14]. Interestingly there are also analogues for morphological scale spaces on matrix fields[17].
Figure 6.2: Top row from left to right: Original tensor field simulating a crossing structure, pseudo-linear with scale 0.2, pseudo-linear with scale $e^{0.2}$, linear with scale 0.2. Bottom row: As in the top row with a noisy original tensor field.

Figure 6.3: Top row from left to right: Original tensor field with the metric tensors of a monkey saddle surface (a parameterized surface embedded in $\mathbb{R}^3$), pseudo-linear with scale 0.2, pseudo-linear with scale $e^{0.2}$, linear with scale 0.2. Bottom row: As in the top row with a noisy original tensor field.
Figure 6.4: Top row from left to right: Original tensor field with various metric tensors to show some effects of blurrings, pseudo-linear with scale 0.2, pseudo-linear with scale $e^{0.2}$, linear with scale 0.2. Bottom row: As in the top row with a noisy original tensor field.

Figure 6.5: Top row from left to right: A planar section of the original three dimensional tensor field with U-shaped bundle, pseudo-linear with scale 0.2, Bottom row: pseudo-linear with scale $e^{0.2}$, linear with scale 0.2.
6.3 A Scale Space for Higher Order Tensors

Since the Stone-Weierstrass theorem guarantees that any function on a sphere can be arbitrarily closely (in $L^2$-norm) approximated by a polynomial, we approximate the diffusion profiles with polynomials. A special feature of HARDI diffusion profiles is that they have axial symmetry. Therefore the orders of the monomials in the polynomial approximation are all even.

Let $F$ be the true diffusion profile. Suppose we have a polynomial approximation $P_N$ of $F$ for some order $N = 2K$. Since only even order monomials are allowed, we can express this polynomial as

$$P_N(x, y) = D(x) + D_{ij}(x) y^i y^j + \cdots + D_{i_1 \cdots i_N}(x) y^{i_1} \cdots y^{i_N}$$

$$= \sum_{n=0}^{N} D_{i_1 \cdots i_N}(x) y^{i_1} \cdots y^{i_N}, \quad (6.44)$$

where

$$y = (y^1, y^2, y^3) = (\sin \theta \cos \varphi, \sin \theta \sin \varphi, \cos \theta). \quad (6.45)$$

We can build this polynomial from the data following [33]. We begin by fitting a zeroth order tensor to data using linear least squares method (3.17) as in Chapters 3 and 5. In case our data was a continuous function $f$ this would amount to taking the integral

$$\frac{1}{4\pi} \int f(\theta, \varphi) \sin \theta d\theta d\varphi. \quad (6.46)$$

The reason that we do not compute the irreducible components simply by an inner product with spherical harmonics is the irregular distribution of data, which would require a partition of the sphere to assign weights to each measurement, before integration. With a least squares approach we obtain the mean of all values as the zeroth order monomial $D_0(x)$. We take the difference

$$E_0(x) = D(x) - D_0(x). \quad (6.47)$$

Next we fit a second order tensor $D_2(x)$ (a polynomial containing only second order monomials) to $E_0(x)$. We subtract this again

$$E_2(x) = D(x) - D_0(x) - D_2(x) = E_0(x) - D_2(x), \quad (6.48)$$

and repeat this procedure up to desired order. As is shown in [33], a tensor $D_i(x)$ obtained this way is equivalent to the linear combination of real $i$th
6.3. A Scale Space for Higher Order Tensors 107

Figure 6.6: Top left: Original profile. Rest: Evolution of this profile with scales 0.01–0.1.

order spherical harmonics, i.e.

\[
\sum_{j=-i}^{i} c_{ij} Y_j^i(\theta, \varphi),
\]

in the spherical harmonic decomposition of \( P_N \). Spherical harmonics are the eigenfunctions of the Laplace-Beltrami operator on the sphere [73]. From this follows that evolution of \( P_N(x, y) \) according to the diffusion equation at time \( \sigma \) equals [16]

\[
P_N(x, y, \sigma) = \sum_{n=0}^{K} e^{-\sigma n(n+1)} D_{i_1 \ldots i_n}(x) y^{i_1} \ldots y^{i_n}. \tag{6.50}
\]

For illustration of Laplace-Beltrami smoothing on a diffusion profile see Fig. 6.6.

In practice \( N \) is always a finite number due to the limited number of directional measurements. Therefore besides being a low-pass filter, this smoothing is also a band-pass filter. As we have shown in Chapter 5 when the maximum number of order (of data approximating tensors) is fixed, we
do not have to actually do the fitting iteratively, but using Clebsch projection, pre-compute the coefficients and smoothing parameters altogether in one matrix. By multiplying the single tensor approximating data with this matrix, we obtain a tensor representing the ODF with scale-parameters at once.

Laplace-Beltrami smoothing is a point wise smoothing method and contributes mostly in dampening the effect of noise. In smoothing one could also take the neighborhood into account as for example in [37].
Conclusions and Future Work

"Tot nog toe lijkt het er sterk op dat het brein van de mens nog niet slim genoeg is om zichzelf te ontdelen en te begrijpen.” Christine Van Broeckhoven
7.1 Conclusions

In this thesis we have studied basic elements of Riemann and Finsler geometry, and applied them to tensor-valued medical images.

Necessary concepts and tools from Riemann geometry were introduced in Chapter 2. The Riemann and the Ricci curvatures were studied in detail and the relation of geodesic deviation and the Ricci curvature were demonstrated. The extremals w.r.t. homogeneous and non-homogeneous Lagrangians and their relation to the Riemannian geodesic equation were also covered in Chapter 2.

In DTI, the successful tracking of axonal bundles is one of the main challenges. The curves obtained using a tractography algorithm do not necessarily correspond to real axonal connection. To assist in selecting the potential fiber candidates, a quality measure for DTI curves was proposed in Chapter 3. Curvature based measures to quantify the local coherence and homogeneity of tensor fields were introduced and the possibility of using the Ricci scalar as an indicator for sudden changes in fiber orientations was studied.

It is well known that modeling diffusion profiles with a second order tensor has serious limitations. To overcome some of these, we extended the geometric framework from Riemannian, which is a natural choice for second order tensor fields, to Finslerian which is suitable for more general tensors. In Chapter 4 the basic elements of Finsler geometry were defined. An alternative geometric formulation of the strong convexity, essential for Finsler geometry, was derived.

In Chapter 5 it was shown that a Finsler framework can be used to study HARDI images. A new way to compute the orientation distribution function efficiently from raw HARDI data was also proposed. This is based on the relation of a single tensor vs. its harmonic decomposition, a subject which has been also explicitly discussed here. In analogy to the DTI case, a quality measure for HARDI-tracts was proposed. We recall that one of the limitations of DTI is that it cannot model crossing bundles. In this chapter, a novel technique for fiber tractography, capable of modeling crossings, was derived with some promising experimental results.

Due to the discrete nature of the data, computations necessarily involve
interpolation. In Chapter 6 the interpolation of tensors was considered and scale-space regularization methods, for both DTI- and HARDI-data, were introduced.

7.2 Future Work

From medical point of view the usefulness of these methods has to be validated. In this thesis we have only seen some positive indications, but further research towards proper interpretations in medical setting is needed. Experiments with high quality phantoms will be essential to justify the use of these, in principle exact methods, on irregular and noisy real data.

From the theoretical aspect there are also many interesting parts that need further studies. The various curvatures in Finsler geometry and their interpretations can be very useful in different areas of geometric data analysis.

The regularization methods studied here could be extended in a direction that takes the neighborhood into account. The first step would be to study Ricci flow on diffusion tensor fields. Extensions of this flow to Finslerian case will require more theoretical studies.

From the implementation point of view an interesting subject is to look for efficient discrete algorithms for differential geometry on tensor fields. Intuitively it is clear that fast template models used in image processing are a reasonable option to begin with. Experiments on new types of curvatures with tensor data could in turn be of inspiration to further theoretical studies. Also a discrete formulation of a connection i.e. covariant derivative on discrete tensor fields could be of interest.

Finally since the methods used in this thesis do not depend on the origin of the tensor data, other types of applications (than medical images) e.g. in mechanics could be fruitful for further research.
Appendix
Some Additional Definitions

We have put here some definitions that either take too much space or interrupt the flow in the main text.

**Definition 8.1.** Let $M$ be a topological $n$-manifold. A coordinate chart is a pair $(U, \varphi)$, where $\varphi : U \rightarrow V$, $U \subset M$ and $V \subset \mathbb{R}^n$. The set $U$ is called a coordinate neighborhood of its points. The homeomorphism (continuous bijection) $\varphi$ is called a local coordinate map. The component functions $\varphi_i$ of $\varphi$ are called local coordinates on $U$. If $(U, \varphi), (V, \psi)$ are two charts s.t. $U \cap V \neq \emptyset$, then the composite map $\psi^{-1} \circ \varphi : \psi(U \cap V) \rightarrow \varphi(U \cap V)$ is called a transition map. An atlas is a collection of charts whose domain cover $M$. An atlas is smooth if the transition maps are smooth (infinitely many times differentiable). A smooth atlas is maximal, if it is not contained in any larger atlas. A smooth manifold is a manifold with smooth atlas. A smooth manifold with symmetric positive definite metric is a Riemannian manifold.

**Definition 8.2.** A coordinate frame on an open subset $U$ of a manifold $M$ is the set of vector fields $(\partial_1, \ldots, \partial_n)$, and this set forms also a basis for $T_pM$ at every $p \in U$.

**Definition 8.3.** A local frame $(E_1, E_2, \ldots, E_n)$ for tangent bundle $TM$ is the set of $n$ smooth vector fields defined on some open set $U$, such that the set $(E_1|_p, \ldots, E_n|_p)$ forms a basis for $T_pM$ at every $p \in U$.

**Definition 8.4.** A vector field $V(t)$ along a curve $\gamma : I \rightarrow M$ is said to be extendible, if there exists a vector field $\tilde{V}$ on $M$, such that $V(t) = \tilde{V}(\gamma(t))$.

In the Figure 8.1, on the left hand side is an extendible vector field and on the right hand side a non-extendible vector field.

**Examples**

The most motivating introduction to differential geometry is given by the classical theory of surfaces [63][56]. For a hands-on introduction containing implementation and visualization we recommend [41]. For an introduction to tensor calculus on manifolds and to Riemannian geometry, we refer the reader to [64][28]. Here we give some examples of Riemannian manifolds.
in dimension two and compute their metric tensors and geodesics for illustration.

**Monkey Saddle**

- **Parametrization**

  \[
  f : \mathbb{R}^2 \to \mathbb{R}^3 : (x, y) \mapsto (x, y, x^3 - 3xy^2). \tag{8.1}
  \]
• Jacobian matrix

\[ J = \begin{pmatrix} \frac{\partial f_1}{\partial x} & \frac{\partial f_1}{\partial y} \\ \frac{\partial f_2}{\partial x} & \frac{\partial f_2}{\partial y} \\ \frac{\partial f_3}{\partial x} & \frac{\partial f_3}{\partial y} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 3x^2 - 3y^2 & -6xy \end{pmatrix} \] (8.2)

• Metric tensor

\[ g = J^T J = \begin{pmatrix} 1 + (3x^2 - 3y^2)^2 & -6xy(3x^2 - 3y^2) \\ -6xy(3x^2 - 3y^2) & 1 + 36x^2y^2. \end{pmatrix} \] (8.3)

• Geodesic equations (2.51), see Fig. 8.2

\[ \ddot{x} = -\frac{1}{c} \left( 18(x^3 - xy^2)\dot{x}\dot{x} + 36y(-x^2 + y^2)\dot{x}\dot{y} + 18(x^3 - xy^2)\dot{y}\dot{y} \right) \]
\[ \ddot{y} = -\frac{1}{c} \left( 36x^2y\dot{x}\dot{x} + 36xy^2\dot{x}\dot{y} + 36x^2y\dot{y}\dot{y} \right), \] (8.4)

where

\[ c = 1 + 9x^4 + 18x^2y^2 + 9y^4. \] (8.5)

Torus

• Parametrization

\[ f : (\theta, \varphi) = ((2 + \cos \theta) \cos \varphi, (2 + \cos \theta) \sin \varphi, \sin \theta). \] (8.6)

• Metric tensor.

\[ \begin{pmatrix} 1 & 0 \\ 0 & (2 + \cos \theta)^2 \end{pmatrix}. \] (8.7)

• Geodesics, see Fig. 8.3

\[ \ddot{\theta}(t) + (2 + \cos \theta \sin \theta)\dot{\varphi}^2(t) = 0 \]
\[ \ddot{\varphi}(t) + \sin \theta(2 + \cos \theta)\dot{\varphi}(t)\dot{\theta}(t) = 0. \]
Sphere

- Parametrization.

\[ f : [0, \pi) \times [0, 2\pi) \to S^2, \quad f : (\theta, \varphi) \mapsto (\sin \theta \cos \varphi, \sin \theta \sin \varphi, \cos \theta) . \]  

(8.8)

- Metric tensor

\[ \begin{pmatrix} 1 & 0 \\ 0 & \sin \theta^2 \end{pmatrix} . \]  

(8.9)

- Geodesics, see Fig. 8.4

\[
\begin{cases}
\ddot{\theta}(t) = \cos \theta \sin \theta \dot{\varphi}^2(t) \\
\ddot{\varphi}(t) = \frac{-2}{\tan \theta} \dot{\varphi}(t) \dot{\theta}(t) .
\end{cases}
\]  

(8.10)
Figure 8.4: Left: geodesics on the parameter plane. Right: geodesics on a sphere embedded in $\mathbb{R}^3$. 
Bibliography


[31] L. Florack and E. Balmashnova. Decomposition of high angular resolution diffusion images into a sum of self-similar polynomials on the sphere. In *Proceedings of the Eighteenth International


[42] B. M. ter Haar Romeny. Front-End Vision and Multi-Scale Image Analysis: Multi-Scale Computer Vision Theory and Applications, written in Mathematica, volume 27 of Computational Imaging and


Summary

Multi-Scale Riemann-Finsler Geometry
Applications to Diffusion Tensor Imaging and High Angular Resolution Diffusion Imaging

In this brief summary, we reflect the goals and the achievements of this PhD-project by citing three sentences in the original project proposal.

"This project investigates the exploitation of the scale degree of freedom in images from the vantage point of differential geometry and tensor calculus in scale space."

Indeed, this thesis introduces the necessary tools for differential geometric tensor calculus and derives some theoretical as well as practical results. The scale is considered separately in two different settings. In this thesis, the weight is somewhat shifted from the problem of scale to problems in differential geometry and applications. The applications considered here are multi-directional medical images, namely diffusion weighted magnetic resonance images. The measurements are modeled using symmetric tensors, which is equivalent to a polynomial approximation. This thesis can be roughly separated into two parts: one that studies second order tensor fields and one that uses higher order tensor fields. The second order tensor fields are studied with tools from Riemann geometry, and the higher order ones respectively with those from Finsler geometry. Each of these can be separated to a theoretical part, that contain the mathematical definitions and derivations and an applied part, where some geometric properties of synthetic, simulated or real data are computed and analyzed.

"The goal is to couple geometry to image content based on a specific task.”

By attaching geometric meaning to the physical properties (of the imaged object) represented by data, we have derived some measures and algorithms to extract information from images. For example a novel method to do fiber tractography, i.e. extract neural connections from diffusion weighted images of brain, is introduced. This method has the special property, that it can propagate through voxels with complex fiber orientations. Techniques to measure relative diffusivity along a curve and to detect inhomogeneities in tensor field are some of the other examples. Since the real data is discrete, the interpolation of tensor fields is also considered.
"The objective is to foster specific applications in biomedical image analysis, and to extend these to multiple scales."

In the Riemannian framework, the concept of scale is introduced in a Gaussian derivative scheme to second order tensor fields. In Finslerian context, a scale parameter was attached to higher order tensors by applying Laplace-Beltrami smoothing, solving the heat equation on the sphere.
Acknowledgements

I am grateful to the members in my defense committee: prof. Bernhard Burgeth, dr. Remco Duits, prof. Jan de Graaf, prof. Luc Florack, prof. Wiro Niessen, prof. Mark Peletier and prof. Bart ter Haar Romeny. I admire Bernhard’s talent to be able to explain the essential even to a non-specialist. Jan has pointed out connections between different areas of analysis, and broadened my point of view. Luc has given my project a motivating start with substantial freedom in my later studies, which I genuinely appreciate. Bart has always been there for his students, even through difficult times, and I have learned a lot from him. I want to thank all my colleagues at the biomedical image analysis group and at the center for analysis, scientific computing and applications (CASA). Especially Evgeniya, I value all of our discussions by the white-board and the numerous useful comments from her. The crowd at the DTI-meetings, that lasted through my whole project (and is still going strong): Anna, Andrei, Bram, Ellen, Erik, Markus, Paulo, Ralph, Remco, Tim, Vesna and others, have greatly contributed to my work. I am also grateful for the participants of the wednesday morning meetings at CASA, especially Adrian, Andryi, Erwin, Georg, Jan Cees, Jim, Mark, Michiel, Kundan and Yabin. Many times have I turned to Bart J., Bas and Ronald, for assistance in scientific computing and appreciate their help. I want to sincerely thank my prestigious paronymfs, Maria and Marieke. Going back in time, it is clear that my project was only possible through the fact that prof. Bob Mattheij generously provided me a spot to work out my Licentiate thesis, which I started at the Helsinki University of Technology. The person who initiated my interest in differential geometry is my Lic. thesis supervisor Kirsi Peltonen, whose rigour and geometric insight I greatly admire. I thank my friends that I have met in Eindhoven: Annikka, Elina, Jaana, Katri, Minna, Nina and Sanna for many "gezellig" occasions. I thank my parents Kristiina and Jaakko for passing on some "Finnish sisu" to me. The cover of this thesis shows the creative chaos, essential for my work, that is generated by our children, Mandi, Juri and Jalo. Last but not least, I want to thank my gorgeous husband Aki Härmä for his unfailing support in everyday life as well as for his professional advices.
Laura Astola was born on December 29th in Turku, Finland. She got her high school diploma from Tottori East prefectural high school in Japan. She began her studies in mathematics at Turku University, but moved later to Helsinki University for family reasons. In April 2000 she finished her Master Thesis on "Computer-aided visualization in mathematics teaching" and began her Ph.D. studies on geometric analysis in Helsinki University of Technology. She completed her Licenciate’s thesis "Lattès type uniformly quasi-regular mappings on compact manifolds" in February 2009, obtaining the degree of Licenciate of Technology. In February 2006 she started as a Ph.D. student within the Biomedical Image Analysis group at Eindhoven University of Technology. In 2007, her supervisor L. Florack was appointed to a professor at the department of Mathematics and Computer Science and she also moved there to become a member of CASA (Center for Analysis, Scientific Computing and Applications).
Publications

L. Florack, E. Balmashanova, L. Astola, E. Brunenberg

A. Fuster, L. Astola, and L. Florack

L. Astola, A. Jalba, E. Balmashnova and L. Florack

L. Astola and L. Florack

L. Astola and L. Florack

L. Astola
Lattès type Uniformly Quasiregular Mappings on Compact Manifolds. Licentiate Thesis, Helsinki University of Technology, Department of Mathematics and Systems Analysis, February, 2009, Finland.

L. Astola, M. van Almsick and L. Florack
L. Astola and L. Florack

L. Florack and L. Astola

L. Astola and L. Florack and B. ter Haar Romeny

L. Astola and L. Florack and B. ter Haar Romeny