1. Een kwadratisch integreerbare afgeleide van een kwadratisch integreerbare functie is een *analyzing wavelet*.
   - Dit proefschrift, hoofdstuk 2.

2. De Haar basis is de enige wavelet basis van $l^2$ die QMF-filters met eindige tijdsduur heeft.
   - Dit proefschrift, hoofdstuk 3.

3. Zij gegeven dat de functie $f$ voldoet aan
   \[ f(x) = \int_x^\infty w(t) f(t) \, dt \quad \forall \, x \geq 0 , \]
   waarin $w$ een bekende continue en positieve functie is waarvoor
   \[ \int_0^\infty w(t) \, dt = \infty \]
   en waarin de functie $p$ gegeven is door
   \[ p(x) \int_x^\infty w(t) \, dt = 1 \quad \forall \, x \geq 0 . \]

   Indien bovendien gegeven is dat $f$ een maximum heeft, of een minimum heeft, dan wel dat $\lim_{x \to \infty} f(x)$ bestaat dan is $f$ constant voor $x \geq 0$.

4. Zij $X(t)$ een stationair, m.s. differentieerbaar, Gaussisch proces met gemiddelde 0 en correlatiefunctie $\Gamma(\tau)$ en zij
   \[ \sigma_1^2 = \Gamma(\tau) \bigg|_{\tau=0} , \quad \sigma_2^2 = -\Gamma'(\tau) \bigg|_{\tau=0} . \]
   Het verwachte aantal malen dat $X(t)$ een differentieerbare functie $h(t)$ doorsnijdt per tijdseenheid is gegeven door
   \[ \frac{\sigma_2}{\pi \sigma_1} e^{-\frac{1}{2} \left( \frac{\hat{h}^2(t)}{\sigma_1^2} + \frac{\hat{h}^2(t)}{\sigma_2^2} \right)} + \frac{\hat{h}(t)}{2\pi \sigma_1 \sigma_2} \int_{-\hat{h}(t)}^{\hat{h}(t)} e^{-\frac{1}{2} \left( \frac{y}{\sigma_2} \right)^2} \, dy . \]

Stellingen

bij het proefschrift

The multiscale and wavelet transform with applications in well log analysis

van

P.L. Vermeer
5. De stelling dat de militaire dienstplicht noodzakelijk is, omdat het leger een afspiegeling dient te zijn van de samenleving, is onhoudbaar.

6. De Eerste Kamer zou direct en volgens een distriktenstelsel moeten worden gekozen.

7. Bij bridge is het belangrijk de tegenstanders niet zomaar een contract op laag niveau te laten spelen. In dit gevecht om de deelscore wordt de ongewone SA te weinig gebruikt.

<table>
<thead>
<tr>
<th>Noord</th>
<th>Oost</th>
<th>Zuid</th>
<th>West</th>
</tr>
</thead>
<tbody>
<tr>
<td>♠ 1</td>
<td>♠ pas</td>
<td>♠ 2</td>
<td>♠ pas</td>
</tr>
<tr>
<td>♠ pas</td>
<td>2SA</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Het 2SA bod van West belooft nu ♠/♦ en circa 8 punten. Hij heeft bijvoorbeeld: ♠ 5 4
♥ 8 5
♦ H V B 3
♣ V B 6 4 2

Een doublet van West zou juist een hand met ♠ samen met ♠ of ♦ beloven.

8. De term malt bier, die door twee grote Nederlandse biermerken wordt gebruikt voor hun alcoholvrij bier, is een pleonasme en geeft geen produktinformatie.

9. Een interessant geval van positieve discriminatie is het gebruik van uitsluitend vrouwelijke hop voor de bereiding van bier.

10. Het verdient aanbeveling om voor de reclameboodschappen op de rechterzijde van bussen en trams een taal te gebruiken met van rechts naar links lezend schrift. Arabisch is de meest voor de hand liggende keuze.

11. Niets is zo frustrerend als in een file staan die niet wordt vermeld op de radio.
The multiscale and wavelet transform
with applications in well log analysis
The multiscale and wavelet transform
with applications in well log analysis

Proefschrift

ter verkrijging van de graad van doctor
aan de Technische Universiteit Delft,
op gezag van de Rector Magnificus,
prof. drs. P.A. Schenck,
in het openbaar te verdedigen
ten overstaan van een commissie
aangewezen door het College van Dekanen
op donderdag 16 januari 1992 te 14.00 uur

door

Pieter Leonard Vermeer,

geboren te Vlaardingen,
wiskundig ingenieur.
Dit proefschrift is goedgekeurd door de promotor:

Prof. dr. ir. A.J. Hermans

Promotiecommissie:

Dr. ir. J.A.H. Alkemade
Prof. dr. ir. E. Backer
Prof. dr. ir. C.J. van Duijn
Prof. dr. ir. A.J. Hermans
Prof. dr. ir. M.A. Viergever
Prof. A.M. Ziolkowski, Ph.D.
Contents

Definitions

1 Introduction

2 Transforms
   2.1 The affine transform
       2.1.1 Definition
       2.1.2 Nature of the affine transform
       2.1.3 Inversion and example
   2.2 The multiscale transform
   2.3 The wavelet transform
       2.3.1 Definition and examples
       2.3.2 Sampling and expansion

3 Discrete orthonormal wavelets
   3.1 Introduction
   3.2 Subband filterbanks and orthonormal wavelet bases
   3.3 Multiresolution representation for $l^2$
       3.3.1 Preliminary
       3.3.2 Construction
   3.4 Wavelet representation for $l^2$
       3.4.1 Preliminary
       3.4.2 Construction
       3.4.3 Reconstruction
   3.5 General formula for the QMF-filters
   3.6 Finite duration QMF-filters
   3.7 Continuity of the scaling and wavelet function
       3.7.1 Decay rate
       3.7.2 Consequence
   3.8 Construction of an example
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.9</td>
<td>Conclusions</td>
<td>40</td>
</tr>
<tr>
<td>3.A</td>
<td>Proof of Theorem 1, 2 and 3</td>
<td>45</td>
</tr>
<tr>
<td>3.A.1</td>
<td>Proof of Theorem 1</td>
<td>45</td>
</tr>
<tr>
<td>3.A.2</td>
<td>Proof of Theorem 2</td>
<td>50</td>
</tr>
<tr>
<td>3.A.3</td>
<td>Proof of Theorem 3</td>
<td>51</td>
</tr>
<tr>
<td>3.B</td>
<td>Proof of Theorem 4</td>
<td>54</td>
</tr>
<tr>
<td>3.C</td>
<td>Miscellaneous</td>
<td>58</td>
</tr>
<tr>
<td>4</td>
<td>Multiscale segmentation of well logs</td>
<td>61</td>
</tr>
<tr>
<td>4.1</td>
<td>Introduction</td>
<td>61</td>
</tr>
<tr>
<td>4.2</td>
<td>Multiscale edge detection</td>
<td>62</td>
</tr>
<tr>
<td>4.2.1</td>
<td>General</td>
<td>62</td>
</tr>
<tr>
<td>4.2.2</td>
<td>Scale space</td>
<td>62</td>
</tr>
<tr>
<td>4.2.3</td>
<td>Numerical aspects</td>
<td>67</td>
</tr>
<tr>
<td>4.3</td>
<td>Multiscale segmentation</td>
<td>77</td>
</tr>
<tr>
<td>4.3.1</td>
<td>General</td>
<td>77</td>
</tr>
<tr>
<td>4.3.2</td>
<td>Modified scale space</td>
<td>77</td>
</tr>
<tr>
<td>4.3.3</td>
<td>Multiscale blocking</td>
<td>78</td>
</tr>
<tr>
<td>4.3.4</td>
<td>Multiscale trend detection</td>
<td>79</td>
</tr>
<tr>
<td>4.3.5</td>
<td>Local descent to finer scales</td>
<td>79</td>
</tr>
<tr>
<td>4.3.6</td>
<td>Comparison with a common method</td>
<td>80</td>
</tr>
<tr>
<td>4.4</td>
<td>Conclusions</td>
<td>82</td>
</tr>
<tr>
<td>Figures section 4.3</td>
<td></td>
<td>83</td>
</tr>
<tr>
<td>5</td>
<td>Signal matching by dynamic programming</td>
<td>89</td>
</tr>
<tr>
<td>5.1</td>
<td>Introduction</td>
<td>89</td>
</tr>
<tr>
<td>5.2</td>
<td>Preliminaries</td>
<td>90</td>
</tr>
<tr>
<td>5.2.1</td>
<td>Problem definition</td>
<td>90</td>
</tr>
<tr>
<td>5.2.2</td>
<td>Parametrization-independent measure</td>
<td>91</td>
</tr>
<tr>
<td>5.2.3</td>
<td>Parametrization</td>
<td>92</td>
</tr>
<tr>
<td>5.2.4</td>
<td>Uniqueness</td>
<td>93</td>
</tr>
<tr>
<td>5.3</td>
<td>Distance measures</td>
<td>98</td>
</tr>
<tr>
<td>5.3.1</td>
<td>General</td>
<td>98</td>
</tr>
<tr>
<td>5.3.2</td>
<td>Measure for equality</td>
<td>98</td>
</tr>
<tr>
<td>5.3.3</td>
<td>Measure for equality with additive variation</td>
<td>101</td>
</tr>
<tr>
<td>5.3.4</td>
<td>Measure for equality with multiplicative variation</td>
<td>103</td>
</tr>
<tr>
<td>5.3.5</td>
<td>Measure for equality with additive and multiplicative variation</td>
<td>107</td>
</tr>
<tr>
<td>5.4</td>
<td>Dynamic Programming</td>
<td>109</td>
</tr>
</tbody>
</table>
5.4.1 General .................................. 109
5.4.2 Formulation as DP-problem .............. 110
5.4.3 Remarks on minimization .................. 112
5.5 Discretization ................................ 113
  5.5.1 The integral expression .................... 113
  5.5.2 The weight function ...................... 115
  5.5.3 Problems for some distance measures ... 116
5.6 Experiments on synthetic data ............... 119
  5.6.1 Preliminary .................................. 119
  5.6.2 Experiment 1 ................................ 121
  5.6.3 Experiment 2 ................................ 126
  5.6.4 Noise .................................. 127
  5.6.5 Additive variation ...................... 129
  5.6.6 Multiplicative variation ................. 132
  5.6.7 Additive and multiplicative variation .. 133
  5.6.8 Deformation with shift .................... 134
  5.6.9 Final remarks ................................ 139
5.7 Application to well logs ..................... 140
  5.7.1 Introduction ................................ 140
  5.7.2 Matching of two well logs ............... 141
  5.7.3 Multiscale evaluation ..................... 144
  5.7.4 Remarks ................................ 148
5.8 Conclusions ................................... 148
Figures Section 5.6 ................................ 151
Figures Section 5.7 ................................ 173

6 Concluding Remarks ............................. 181

References ........................................ 183

Acknowledgement .................................... 187

Summary ............................................ 189

Samenvatting ....................................... 193

Curriculum Vitae .................................. 197
Definitions

$L^p(I\mathbb{R})$ is the space of functions

$$f : I\mathbb{R} \to \mathbb{C},$$

for which

$$\int_{-\infty}^{\infty} |f(x)|^p \, dx < \infty, \quad p \geq 1.$$

The inner product is defined as

$$<f, g> = \int_{-\infty}^{\infty} f(x) g^*(x) \, dx.$$

(* denotes the complex conjugate)

The Fourier transform of a function $f$ is defined by

$$\tilde{f}(\omega) = \int_{-\infty}^{\infty} f(x) e^{-i\omega x} \, dx,$$

with inverse transform

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \tilde{f}(\omega) e^{i\omega x} \, d\omega.$$

$B L^p(I\mathbb{R})$ is the subspace of bounded functions in $L^p(I\mathbb{R})$.

$L^p$ is the space of functions

$$u : \mathbb{Z} \to \mathbb{C}$$
for which
\[ \sum_{n=-\infty}^{\infty} |u(n)|^p < \infty, \quad p \geq 1. \]

The inner product is defined as
\[ (u, v) = \sum_{n=-\infty}^{\infty} u(n)v^*(n). \]

The discrete-time Fourier transform of a function \( u \) is defined by
\[ U(\omega) = \sum_{n=-\infty}^{\infty} u(n)e^{-i\omega n}, \]
with inverse transform
\[ u(n) = \frac{1}{2\pi} \int_{-\pi}^{\pi} U(\omega)e^{i\omega n} \, d\omega. \]

The characteristic function is defined by
\[ \chi_{I}(x) = \begin{cases} 1 & \text{if } x \in I \\ 0 & \text{if } x \notin I \end{cases} \]

The center of a function \( f \) is defined by
\[ x_c(f) = \frac{\int_{-\infty}^{\infty} x |f(x)|^2 \, dx}{\int_{-\infty}^{\infty} |f(x)|^2 \, dx}. \]

The effective width of \( f \) is defined by
\[ \Delta x(f) = \left( \frac{\int_{-\infty}^{\infty} (x - x_c)^2 |f(x)|^2 \, dx}{\int_{-\infty}^{\infty} |f(x)|^2 \, dx} \right)^{\frac{1}{2}}. \]

The effective support of \( f \) is given by
\[ (x_c - \frac{1}{2}\Delta x, x_c + \frac{1}{2}\Delta x). \]
Chapter 1

Introduction

Signal transforms play an important role in the processing of signals and images. The choice of the transform, that will be used for a certain application, depends on the particular properties of a signal or image that one wants to reveal, enhance or suppress.

Probably the best-known transform in signal processing is the Fourier transform. The Fourier transform of a signal yields an expansion in terms of sines and cosines. Since these functions are of infinite duration, the time (spatial) behavior of a signal is concealed in the Fourier or frequency domain. In the frequency domain the global properties of a signal are revealed better, while the local properties are more evident in the time domain. However, since many naturally occurring signals and images have typical global and local behavior, signal transforms are desired that remain, in some sense, in between the time domain and the frequency domain. In this thesis we will distinguish between two interrelated types of transforms that exhibit this particular property.

A multiscale transform results in a description of a signal at different levels of smoothness, usually referred to as scales\(^1\). At a large scale the signal has been smoothed strongly, thus leaving only the global properties. A small scale means less smoothing, thus maintaining more of the local properties. A well-known multiscale transform is the Gaussian pyramid. It consists of a sequence of signals each of which is the smoothed and down-sampled version of its predecessor. Since down-sampling reduces the length of a finite signal, the result is a sequence of signals of decreasing length, which explains the name pyramid. The prefix Gaussian is derived from a particular choice for the applied smoothing filter.

A time-frequency transform results in a description of the frequency content

\(^1\)In literature scale is often called resolution, in which case the multiscale transform is called multiresolution transform, representation or decomposition.
of a signal as a function of time. The signal is thereby filtered with functions that are localized both in time and in frequency. As a consequence of the uncertainty principle for signals (Gabor[14]), the frequency content of a signal at a certain time can only be given with finite accuracy. Distinction can be made between two types of time-frequency transforms:

- The first type is the so-called short-time Fourier transform. In this transform the Fourier transform is applied to a part of the signal within a window, that shifts along the entire length of the signal. The transform performs a filter operation on the signal with functions of which the amplitude is exactly the shifted window function while their phase is the complex exponential in the Fourier transform. Time is controlled by the shift of the window, while frequency is controlled by modulation. It is noted that several possibilities exist for the window function. If for instance the window function is the Gaussian the transform is known as the Gabor transform.

- The second type of time-frequency transform is the wavelet transform. This transform filters the signal with a set of functions, which are obtained by shifts and dilations of a certain basic function. Hence, time is again controlled by shifts but frequency is controlled by dilation. In the past few years a considerable amount of wavelet theory has been developed. This includes theory on inverse transformation, sampling and expansion.

In this thesis the attention is focussed on theory and application of the multiscale transform and the wavelet transform.

As for the applications, most attention will be given to the analysis of well logs. Well log analysis is a subject that stems from geological and geophysical data processing. A well log is a sequence of measurements that is obtained from a bore hole. A measuring device, the logging tool, which is attached to a cable, is lowered in the bore hole. Approximately every half foot the logging tool measures one or more physical quantities. Examples are the natural gamma radiation, the spontaneous potential, the sound velocity and the specific density.

The well logs are a record of the variation in lithology. Important in the analysis of well logs is the recognition of boundaries of geological units. The subdivision of well logs in geological units is known as segmentation or zoning.
As part of a geophysical survey, well logs are taken from different bore holes in a certain area. The well logs may exhibit differences due to different geological circumstances at the various locations. Dissimilar rates of sedimentation will cause the well logs to be stretched or compressed with respect to each other. Erosion or faulting may cause a part of a well log to be truncated with respect to the other well logs. Well log matching or correlation is the determination of the similarity between the well logs, in order to obtain a geophysical cross section of the subsurface.

The organization of this thesis is as follows. Preceding this introduction a list of definitions has been included. In the text they will be neither defined nor referenced.

In Chapter 2 the multiscale transform and the wavelet transform are defined. Theory concerning inversion and sampling are reviewed. Some properties and examples are given.

Chapter 3 is concerned with the theory of discrete orthonormal wavelet bases for discrete-time signals. Necessary properties of these wavelets are considered. It will appear that the expansion of discrete-time signals in terms of discrete orthonormal wavelets can be calculated efficiently using an exact-reconstruction filterbank.

Examination of well logs leads to the observation that they exhibit characteristic behavior over a wide range of scales, differing from one foot to hundreds of feet. This behavior is not accounted for in conventional segmentation methods. In Chapter 4 of this thesis a segmentation method is proposed to solve this problem. The segmentation method is based on the multiscale transform and the wavelet transform of the well log. Examples as well as possibilities for further application of the segmentation are provided.

In Chapter 5 a signal-matching method is derived and examined that can determine the shifts, stretches and compressions of which the deformation between two well logs may consist. The method involves a minimization technique based on dynamic programming. Experiments are carried out on synthetic data as well as on real well logs. Once the deformation has been determined, there is a need to evaluate the matching result. It is necessary to distinguish between segments of one well log that have a clear counterpart in the other
well log and segments that have not. To this purpose the previously mentioned segmentation method will be applied.

Chapter 6 contains the concluding remarks.
Chapter 2

Transforms

2.1 The affine transform

2.1.1 Definition

The affine transform of a function \( f \) is obtained by convolution of \( f \) with dilated versions of a certain analysis filter \( g \),

\[
A(f)(x, \sigma) = c(\sigma) \int_{-\infty}^{\infty} f(t) g\left(\frac{x-t}{\sigma}\right) dt ,
\]

(2.1.1)

for \( \sigma > 0 \).

\( \sigma \) governs the dilation of \( g \) and will be called the scale parameter.

The defining integral exists and is bounded if

\[
f \in L^p(I\mathbb{R}) \text{ and } g \in L^q(I\mathbb{R}) \text{ with } \frac{1}{p} + \frac{1}{q} = 1 .
\]

(2.1.2)

The Fourier transform of eq.(2.1.1) with respect to \( x \) is given by

\[
\tilde{A}(f)(\omega, \sigma) = c(\sigma) \sigma \tilde{f}(\omega) \tilde{g}(\sigma \omega).
\]

(2.1.3)

In the sequel, if explicit conditions are absent, it is tacitly assumed that all functions, products and convolutions can be (inverse-) Fourier transformed (see Butzer and Nessel[5]).

---

2A part of this chapter has appeared in a report of the Faculty of Technical Mathematics and Informatics (Alkemade and Vermeer[1]).
2.1.2 Nature of the affine transform

We will qualitatively examine some of the properties of the affine transform for two kinds of filters. \( g \) can be either a band-pass filter or a low-pass (smoothing) filter. Suppose that \( g \) in eq.\((2.1.1)\) is centered around \( x = x_c \) and has effective width \( \Delta x \) and suppose that \( |\tilde{g}| \) consists of one major lump, which is centered around \( \omega = \omega_c \) and which has effective width \( \Delta \omega \). Dilated and shifted versions of \( g \),

\[
g_x^\sigma(t) = c(\sigma) \ g \left( \frac{x - t}{\sigma} \right),
\]

are consequently centered in time around \( t = x - x_c \sigma \) with effective support

\[
x - \sigma(x_0 + \frac{1}{2}\Delta x) < t < x - \sigma(x_0 - \frac{1}{2}\Delta x).
\]

In the frequency domain the centering of the major lumps is now around \( \omega = \frac{\omega_c}{\sigma} \) with effective support

\[
\frac{\omega_c - \frac{1}{2}\Delta \omega}{\sigma} < \omega < \frac{\omega_c + \frac{1}{2}\Delta \omega}{\sigma}.
\]

Eqs.\((2.1.5)\) and \((2.1.6)\) show which part in time and in frequency of the function \( f \) is analyzed by \( g_x^\sigma \). In eq.\((2.1.5)\) it can be seen that the parameter \( x \) causes the filter to shift along the entire time axis, while the parameter \( \sigma \) alters its effective width. In eq.\((2.1.6)\) the parameter \( \sigma \) causes the filter to shift along the frequency axis whilst changing the effective bandwidth.

If \( \omega_c = 0 \) only the effective bandwidth changes as \( \sigma \) changes. The affine transform therefore performs either band-pass or low-pass filter operations on the function \( f \) within a certain window determined by eq.\((2.1.5)\).

If \( g \) is a low-pass filter then \( A(f) \) constitutes a multiscale or multiresolution representation of \( f \). The affine transform then will be called the multiscale transform (Section 2.2).

In the case \( g \) is a band-pass filter, \( A(f) \) constitutes a time-frequency representation of \( f \). Under certain additional conditions the affine transform then becomes the wavelet transform (Section 2.3).
2.1.3 Inversion and example

In this section an inversion formula for the affine transform will be derived that takes into account the entire \((x, \sigma)\)-plane. Another inversion formula, which involves only one \(\sigma\)-value, is discussed in Section 2.2.

Consider eq.(2.1.3),

\[
\tilde{A}(f)(\omega, \sigma) = c(\sigma) \sigma \tilde{f}(\omega) \tilde{g}(\sigma \omega) .
\]

Suppose we have a certain synthesis filter \(\tilde{h}(\sigma \omega)\). Multiplication of both sides of this equation with \(\tilde{h}(\sigma \omega)\) and division of both sides by \(c(\sigma)\sigma^2\) results in

\[
\frac{\tilde{h}(\sigma \omega) \tilde{A}(f)(\omega, \sigma)}{c(\sigma)\sigma^2} = \frac{1}{\sigma} \tilde{g}(\sigma \omega) \tilde{h}(\sigma \omega) \tilde{f}(\omega) .
\]

Hence,

\[
\int_{0}^{\infty} \frac{\tilde{h}(\sigma \omega) \tilde{A}(f)(\omega, \sigma)}{c(\sigma)\sigma^2} \, d\sigma = \tilde{f}(\omega) \int_{0}^{\infty} \tilde{g}(\sigma \omega) \tilde{h}(\sigma \omega) \frac{d\sigma}{\sigma} .
\]

Substitution of \(\nu = \sigma \omega\) in the right-hand-side integral gives

\[
\int_{0}^{\infty} \frac{\tilde{h}(\sigma \omega) \tilde{A}(f)(\omega, \sigma)}{c(\sigma)\sigma^2} \, d\sigma = \begin{cases} 
\tilde{f}(\omega) \int_{0}^{\infty} \tilde{g}(\nu)\tilde{h}(\nu) \frac{d\nu}{\nu} , & \omega > 0 , \\
\tilde{f}(\omega) \int_{0}^{\infty} \tilde{g}(-\nu)\tilde{h}(-\nu) \frac{d\nu}{\nu} , & \omega < 0 ,
\end{cases}
\tag{2.1.7}
\]

Thus, if an analysis/synthesis filter pair \(g, h\) fulfills the conditions

\[
c_{gh}^+ = \int_{0}^{\infty} \tilde{g}(\omega)\tilde{h}(\omega) \frac{d\omega}{\omega} < \infty ,
\tag{2.1.8}
\]

\[
c_{gh}^- = \int_{0}^{\infty} \tilde{g}(-\omega)\tilde{h}(-\omega) \frac{d\omega}{\omega} < \infty ,
\tag{2.1.9}
\]

the Fourier transform of \(f\) can be recovered from its affine transform.
In the special case that \( c_{gh} = c_{gh}^+ = c_{gh}^- \), the inversion formula becomes

\[
f(x) = \frac{1}{c_{gh}} \int_0^\infty \int_{-\infty}^\infty \frac{h(x-t)}{c(\sigma)\sigma^3} A(f)(t, \sigma) \, dt \, d\sigma.
\] (2.1.10)

Eqs.(2.1.8) and (2.1.9) are called admissibility conditions for the filter pair \( g, h \).

**Example**

The filter pair

\[
\begin{align*}
g(x) &= \frac{1}{\sqrt{2\pi}} \frac{d^2}{dx^2} e^{-\frac{1}{2}x^2} ; \quad \widetilde{g}(\omega) = -\omega^2 e^{-\frac{1}{2}\omega^2}, \\
h(x) &= \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} ; \quad \widetilde{h}(\omega) = e^{-\frac{1}{2}\omega^2}
\end{align*}
\]

is an admissible analysis/synthesis filter pair for the affine transform, which, in this case, constitutes a time-frequency representation. If \( g \) and \( h \) are interchanged, the affine transform yields a multiscale representation.

### 2.2 The multiscale transform

If the filter \( g \) in eq.(2.1.1) is a low-pass or smoothing filter the affine transform becomes the multiscale transform. The multiscale transform mainly serves analysis purposes and is not used for filtering in the affine-transform domain. It is therefore that the inversion formula of the previous section will not be used for the multiscale transform.

The condition usually imposed on the multiscale transform is that it should tend to \( f \) for decreasing scale \( \sigma \),

\[
\lim_{\sigma \to 0} A(f)(x, \sigma) = f(x)
\] (2.2.1)

(Witkin[39], Babaud et al.[2], Yuille and Poggio[41]).

Theory on the subject of limit behavior of convolutions, with the type of filters we have, can be found in Butzer and Nessel[5]. Here we will only give a brief summary.
The property in eq.(2.2.1) can be ensured if the following three conditions are met

C.1 \[ g \in L^1(\mathbb{R}) \]

C.2 \[ c(\sigma) \int_{-\infty}^{\infty} g\left(\frac{x}{\sigma}\right) \, dx = 1 \quad \forall \sigma > 0 . \]

A consequence is that

\[ c(\sigma) = \frac{1}{\sigma} , \]

C.3 \[ \lim_{\sigma \to 0} \frac{1}{\sigma} \int_{|x| \geq \delta} |g\left(\frac{x}{\sigma}\right)| \, dx = 0 \quad \forall \delta > 0 . \]

In Butzer and Nessel[5] the affine transform is known as a singular integral and the set of filters \( \left\{ \frac{1}{\sigma} g\left(\frac{x}{\sigma}\right), \sigma > 0 \right\} \) is called an approximate identity. The limit in eq.(2.2.1) can exist in various senses, for our purpose it is sufficient to know that pointwise convergence is assured in each point of continuity of \( f \).

If we look at the problem less strictly, it can readily be seen in eq.(2.1.3), that for eq.(2.2.1) to hold, we should have

\[ c(\sigma) = \frac{1}{\sigma \tilde{g}(0)} \]

\[ = \frac{1}{\sigma \int_{-\infty}^{\infty} g(x) \, dx} , \]

provided

\[ \tilde{g}(0) = \int_{-\infty}^{\infty} g(x) \, dx \neq 0 . \]

The smoothing filter should tend to the delta function as \( \sigma \) tends to 0,

\[ \lim_{\sigma \to 0} \frac{1}{\sigma} g\left(\frac{x}{\sigma}\right) = \delta(x) . \]
Definition

The multiscale transform of a function $f$ is defined by

$$M(f)(x, \sigma) = \frac{1}{\sigma} \int_{-\infty}^{\infty} f(t) g\left(\frac{x-t}{\sigma}\right) \, dt,$$

(2.2.2)

where $g$ is symmetric and satisfies conditions C.1, C.2 and C.3.

$g$ symmetric guarantees that the multiscale transform of a symmetric function is also symmetric. The consequence is that the odd derivatives of $g$ all vanish in the origin. Hence,

$$g^{(n)}(0) = 0 \text{ for } n \text{ is odd and positive.}$$

Here we have adopted the notation: $g^{(n)}(a) = \frac{d^n}{dx^n} g(x) \bigg|_{x=a}$.

Example

In fig.2.2.1 the multiscale transform of a signal is depicted at three scale levels. The applied filter is the Gaussian

$$g(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}.$$

(2.2.3)

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure2_2_1.png}
\caption{Multiscale transform of a signal for three $\sigma$ values}
\end{figure}
2.3 The wavelet transform

2.3.1 Definition and examples

In Morlet et al.[25],[26] the wavelet transform was introduced as a new method to obtain a time-frequency representation of seismic signals. However, in 1967 Speiser[29] defined the wide-band cross ambiguity function, which is precisely the same.

Definition
The wavelet transform of a function $f$ is defined by

$$W(f)(x, \sigma) = \frac{1}{\sqrt{\sigma}} \int_{-\infty}^{\infty} f(t) \psi^*(\frac{t-x}{\sigma}) \, dt,$$

where $\psi$ has to satisfy the admissibility conditions

$$c^+_{\psi} = \int_{0}^{\infty} \frac{|\tilde{\psi}(\omega)|^2}{\omega} \, d\omega < \infty,$$  \hspace{1cm} (2.3.1)

$$c^-_{\psi} = \int_{0}^{\infty} \frac{|\tilde{\psi}(-\omega)|^2}{\omega} \, d\omega < \infty.$$  \hspace{1cm} (2.3.2)

The function $\psi$ is called the analyzing wavelet. The shifted and dilated versions of $\psi$,

$$\psi^\sigma_x(t) = \frac{1}{\sqrt{\sigma}} \psi(\frac{t-x}{\sigma}),$$

are called wavelets. The wavelet transform is a special case of the affine transform (eq.(2.1.1)) with $g(x) = \psi^*(-x)$, where the analysis and synthesis filters are connected by

$$g(x) = h^*(-x),$$

or equivalently $\tilde{g}(\omega) = \tilde{h}^*(\omega)$,

and in which $c(\sigma)$ has been chosen such that all wavelets have equal energy

$$\int_{-\infty}^{\infty} |\psi^\sigma_x(t)|^2 \, dt = \int_{-\infty}^{\infty} |\psi(t)|^2 \, dt.$$
Note that the admissibility conditions of eqs. (2.3.1) and (2.3.2) coincide if the analyzing wavelet $\psi$ is real valued, since in that case $\tilde{\psi}(\omega) = \psi^*(-\omega)$.

A square-integrable derivative of any square-integrable function is an analyzing wavelet. This becomes clear if we realize that the Fourier transform of the derivative of some function $\phi$ is given by

$$i\omega \tilde{\phi}(\omega).$$

Substitution in the admissibility condition in eq. (2.3.1) leads to

$$c^+_\psi = \int_0^\infty \omega |\tilde{\phi}(\omega)|^2 \, d\omega$$

$$\leq \int_0^1 |\tilde{\phi}(\omega)|^2 \, d\omega + \int_1^\infty \omega^2 |\tilde{\phi}(\omega)|^2 \, d\omega \leq ||\phi||_2^2 + ||\phi'||_2^2$$

and for $c^-_\psi$ in eq. (2.3.2) a similar result can be obtained. Parseval's theorem and the square integrability of both $\phi$ and its derivative now lead to the above assertion.

In wavelet theory the attention is often restricted to wavelet transforms of so-called Hardy functions, which are functions $f \in L^2(\mathbb{R})$ for which

$$\tilde{f}(\omega) = 0, \quad \omega < 0$$

(see Grossmann and Morlet[16]). In that case only the admissibility condition of eq. (2.3.1) is relevant, as can be seen in eq. (2.1.7).

**Examples**

1. The Haar function

$$\psi(x) = -\chi_{[0,\frac{1}{2})}(x) + \chi_{[\frac{1}{2},1)}(x),$$

with Fourier transform

$$\tilde{\psi}(\omega) = \frac{4}{i\omega} e^{-\frac{1}{2}i\omega} \sin^2(\frac{\omega}{4}).$$

is an analyzing wavelet.
2. The function

\[ \psi(x) = \frac{1}{\sqrt{2\pi}} \frac{d^2}{dx^2} e^{-\frac{1}{2}x^2}, \]

with Fourier transform

\[ \tilde{\psi}(\omega) = -\omega^2 e^{-\frac{1}{2}\omega^2}, \]

is an admissible analysis/synthesis filter for the affine transform and an analyzing wavelet for the wavelet transform.

In fig.2.3.1 the wavelet transform of a signal is depicted for three \( \sigma \) values.

![Wavelet transform of a signal for three \( \sigma \) values](image)

Figure 2.3.1: Wavelet transform of a signal for three \( \sigma \) values
2.3.2 Sampling and expansion

The inversion formula of eq.(2.1.10) can be read in two ways. On the one hand it shows how to reconstruct the original function from its transform. On the other hand it shows how a function can be written as a superposition of the basic functions \( h\left(\frac{x-t}{\sigma}\right) \).

For the discrete case two problems are of importance. Firstly, the reconstruction of the original function from its sampled wavelet transform. Secondly, the expansion of a function in a discrete set of wavelets.

Suppose the wavelet transform of \( f \) is sampled in both time and scale as follows

\[
W(f)(n\tau \sigma^m, \sigma^m) = \langle f, \psi_n^m \rangle \quad \forall \ n, m \in \mathbb{Z},
\]

(2.3.3)

in which \( \sigma \) and \( \tau \) are constant sampling parameters and where

\[
\psi_n^m(t) = \frac{1}{\sigma^{\frac{1}{2}m}} \psi\left(\frac{t - n\tau \sigma^m}{\sigma^m}\right),
\]

(2.3.4)

\[
= \frac{1}{\sigma^{\frac{1}{2}m}} \psi(\sigma^{-m}t - n\tau) \quad \forall \ n, m \in \mathbb{Z}.
\]

If we recall that the relation between \( \sigma \) and frequency is expressed in eq.(2.1.6) by \( 1/\sigma \sim \omega \), the sampling distance in time in eq.(2.3.4) can be seen to decrease exponentially with increasing frequency, or

\[
\Delta = \tau \sigma^m \sim \frac{\tau}{\omega^m}.
\]

The reconstruction problem is in formula to find functions \( \xi_n^m \), with \( n, m \in \mathbb{Z} \), such that

\[
f = \sum_{n,m} \langle f, \psi_n^m \rangle \xi_n^m.
\]

(2.3.5)

The expansion problem is in formula to find functions \( \xi_n^m, n, m \in \mathbb{Z} \), such that

\[
f = \sum_{n,m} \langle f, \xi_n^m \rangle \psi_n^m.
\]

(2.3.6)

If the set \( \{\psi_n^m\}_{n,m} \) is complete and orthonormal the \( \xi_n^m \)'s are simply the \( \psi_n^m \) themselves. There are however more general possibilities of which we will give an example but first we define the concept of a frame.
A set of vectors \( \{e_i\}_{i \in I} \) in a Hilbert space \( V \) is a frame if there exist constants \( 0 < A \leq B < \infty \) such that for all \( f \in V \)

\[
A \left( f, f \right) \leq \sum_{i \in I} |(f,e_i)|^2 \leq B \left( f, f \right).
\]

The constants \( A \) and \( B \) are called the frame bounds. A frame is a complete set of vectors (Young[40]). It can be shown that there exists a set of vectors \( \{d_i\}_{i \in I} \), which itself is a frame, such that

\[
f = \sum_{i \in I} (f, d_i) e_i = \sum_{i \in I} (f, e_i) d_i .
\]  \hfill (2.3.7)

If the set of vectors \( \{e_i\}_{i \in I} \) is independent there even exists a unique bi-orthonormal set of vectors \( \{d_i\}_{i \in I} \). This is a set of vectors that satisfies the previous equations and for which moreover

\[
(d_i, e_j) = \begin{cases} 
1 & \text{if } i = j , \\
0 & \text{if } i \neq j .
\end{cases}
\]

In the special case that \( A \) and \( B \) coincide \( \{e_i\}_{i \in I} \) is called a tight frame. Eq.(2.3.7) then takes the simple form

\[
f = \frac{1}{A} \sum_{i \in I} (f, e_i) e_i
\]

Note that a complete orthogonal set is a special kind of tight frame.

Whether the \( \ell_n^m \)'s themselves are wavelets, in the case that the \( \varphi_n^m \)'s only constitute a frame and not a tight frame, remains a conjecture.

An extensive discussion on the construction of frames and tight frames of wavelets can be found in Daubechies et al.[8] and Daubechies[9].

Examples

1. Daubechies[9] has shown that with the analyzing wavelet

\[
\psi(x) = \frac{2}{\sqrt{3}} \left( \frac{1}{\pi} \right)^{\frac{1}{4}} (1 - x^2) e^{-\frac{1}{2}x^2}
\]
and with \( \sigma = 1.5 \) and \( \tau = 1 \) in eq.(2.3.3), the set \( \{ \psi_n^m \}_{n,m \in \mathbb{Z}} \) is a frame with frame bounds \( A = 5.79 \) and \( B = 5.86 \).

2. With \( \psi \) the Haar function

\[
\psi(x) = -\chi_{[0,\frac{1}{2})}(x) + \chi_{[\frac{1}{2},1)}(x),
\]

and with \( \sigma = 2 \) and \( \tau = 1 \) in eq.(2.3.3), the set \( \{ \psi_n^m \}_{n,m \in \mathbb{Z}} \) constitutes an orthonormal basis. This basis is usually called the Haar basis.
Chapter 3

Discrete orthonormal wavelets

3.1 Introduction

Multiresolution representations of signals receive considerable attention in signal and image processing. A multiresolution representation describes a signal or image at different levels of smoothness, usually referred to as scales. This approach is useful in signal and image analysis because it provides a separation between local and global behavior.

A subband representation is a decomposition of a signal in frequency bands. One subband can be seen as the difference between two adjacent scales in the multiresolution representation. A subband representation is consequently less redundant than a multiresolution representation. Preferably, the signal represented in one subband is orthogonal to the signal represented in the other subbands. Within most of the subbands the signal appears to be largely decorrelated and it is therefore that this procedure is applied extensively in data compression.

With respect to the multiresolution representation of a signal, we are interested in:

- the filter that connects two scales.
- the filter that connects a certain scale to the original signal.

As for the subband representation, we look for:

- the filter that connects a scale to a subband.
- the filter that connects a certain subband to the original signal.

These problems were raised in Burt and Adelson[4] and they were answered to great extent in Mallat[22]. The theory presented in the latter provides the

This chapter has appeared as a report of the Faculty of Technical Mathematics and Informatics (Vermeer[34]).
means to build a multiresolution representation of a continuous-time signal. The representation is discrete in scale and at each scale discrete in time. Two adjacent scales are connected with a discrete-time filter. A representation at each scale is connected to the original signal by a continuous-time filter. For each scale this filter is a shifted and/or dilated version of one function, called the scaling function. From this scaling function, a so-called wavelet function can be obtained. Filtering the original signal with shifted and dilated versions of this function, which are called wavelets, results in its decomposition in sub-bands. The important conditions that are imposed on the wavelets are that they are orthonormal and that almost any continuous-time signal can be expanded in a series of wavelets.

The problem we are concerned with in this chapter is whether it is possible to develop a theory for the construction of a multiresolution and subband representation for discrete-time signals. This means that the filter that connects a certain scale or subband to the original signal should be discrete-time as well. In the end this will result in the construction of discrete orthonormal wavelets. This subject is also given attention in Evangelista[13]. In the next section we will give a review of multiresolution representations in more detail with emphasis on the theory of Mallat[22]. In Sections 3.3 and 3.4 we derive the basic theory and formulas. The remaining sections are concerned with the construction of filters with the properties described above.
3.2 Subband filterbanks and orthonormal wavelet bases

The review given here is obtained for the greater part from Alkemade and Vermeer[1]. A subband filterbank decomposes a signal in frequency bands (Vaidyanathan[30]). Its concept is illustrated in the scheme depicted in fig.3.2.1. The signal $s$ is fed into the analysis part of the filter bank that consists of a low-pass (smoothing) filter $h_A$ and a high-pass filter $g_A$. After the filtering the results are sampled down. In formula the relation between adjacent levels is given by

$$s_j(n) = \sum_{k=-\infty}^{\infty} h_A(2n - k) \ s_{j-1}(k) ,$$  \hspace{1cm} (3.2.1)

$$r_j(n) = \sum_{k=-\infty}^{\infty} g_A(2n - k) \ s_{j-1}(k) ,$$  \hspace{1cm} (3.2.2)

for $j \geq 1$ and with $s_0 = s$. 

$r_j$ represents the signal in subband $j$. $s_j$ represents the signal at resolution or scale $j$ and can be used for further subband decomposition. We are interested in the filters $w_j$ and $u_j$ that satisfy

$$s_j(n) = \sum_{k=-\infty}^{\infty} w_j(2^j n - k) s(k), \quad (3.2.3)$$

$$r_j(n) = \sum_{k=-\infty}^{\infty} u_j(2^j n - k) s(k), \quad (3.2.4)$$

for $j \geq 1$ and with $s_0 = s$.

As has been noted in the introduction this question was raised in Burt and Adelson[4], in which only numerical calculations for $w_j$ given $h_A$ have been performed. The filter operations in eqs. (3.2.3) and (3.2.4) involve a convolution and subsequent downsampling with a factor $2^j$. The filter $w_j$ in eq. (3.2.3) is a low-pass (smoothing) filter and the filter $u_j$ in eq. (3.2.4) is a band-pass filter.

Exact reconstruction of a signal from its subbands is possible if the filters $h_A$, $g_A$, $h_R$ and $g_R$ satisfy certain conditions. Exact reconstruction is guaranteed if

$$|H_A(\omega)|^2 + |H_A(\omega + \pi)|^2 = 1,$$

$$H_R(\omega) = H_A^*(\omega),$$

$$G_A(\omega) = e^{i\omega} H_R(\omega + \pi),$$

$$G_R(\omega) = G_A^*(\omega).$$

The filters $H_A$ and $G_A$ are called Quadrature Mirror Filters (QMF).

The theory of Mallat[22] contains formulas that resemble eqs. (3.2.3) and (3.2.4) for exact reconstruction filter banks. Since we do not want to go into too much detail we will only discuss the equations briefly.

Mallat[22] defines a function $\phi \in L^2(\mathbb{R})$ by

$$\phi(\frac{x}{2}) = 2 \sum_{n=-\infty}^{\infty} h(n) \phi(x - n)$$
with \( h(n) = h_A(-n) \) and a set of functions \( \phi_n^j \) by
\[
\phi_n^j(x) = \sqrt{2^{-j}} \phi(2^{-j} x - n) \quad \text{for} \quad x \in \mathbb{R} \quad \text{and} \quad n, j \in \mathbb{Z} .
\]
It appears that for every \( j \), \( \{\phi_n^j\}_{k=-\infty,\infty} \) is a set of orthonormal functions.

Because the theory of Mallat\[22]\ applies to continuous-time signals, any discrete-time signal \( s \) has to be interpolated first. Let \( f \in L^2(\mathbb{R}) \) be a function such that
\[
s(n) = < f, \phi_n^0 > .
\]
We note that \( f \) is not unique, because if \( f \) is a solution so is any function \( f + g \) for which
\[
< g, \phi_n^0 > = 0 \quad \forall \ n .
\]
The level \( j \) signal \( s_j \) now satisfies
\[
s_j(n) = < f, \phi_n^j > \quad \text{for} \quad j \geq 0 . \tag{3.2.5}
\]
A similar expression exists for \( r_j \). If \( \psi \in L^2(\mathbb{R}) \) is a function that satisfies
\[
\psi\left(\frac{x}{2}\right) = 2 \sum_{n=-\infty}^{\infty} g(n) \phi(x-n)
\]
with \( g(n) = g_A(-n) \) and if the \( \psi_n^j \)’s are defined analogously to the \( \phi_n^j \)’s then \( r_j \) is given by
\[
r_j(n) = < f, \psi_n^j > \quad \text{for} \quad j \geq 1 . \tag{3.2.6}
\]
Eqs.\(3.2.5\) and \(3.2.6\) differ from the objective in eqs.\(3.2.3\) and \(3.2.4\). The \( \psi_k^j \)’s are called wavelets and it can be shown that \( \{\psi_k^j\}_{k=-\infty,\infty} \) constitutes an orthonormal basis of \( L^2(\mathbb{R}) \).

The representation of \( f \) in subband \( j \) is given by
\[
f_j(x) = \sum_{n=-\infty}^{\infty} < f, \psi_n^j > \psi_n^j(x) ,
\]
\[
= \sum_{n=-\infty}^{\infty} r_j(n) \psi_n^j(x) \tag{3.2.7}
\]
and the orthonormality of the \( \psi_k^j \)’s leads to
\[ <f_j, f_k> = \begin{cases} 1 & \text{if } j = k, \\ 0 & \text{if } j \neq k. \end{cases} \tag{3.2.8} \]

The theory presented in this chapter is the discrete analogue of the theory of Mallat\cite{mallat1989} and results in equations of the form of eqs.\( (3.2.3) \), \( (3.2.4) \). The orthogonality of the subbands appears in discrete equivalents of eqs.\( (3.2.7) \) and \( (3.2.8) \).

\section{Multiresolution representation for \( l^2 \)}

\subsection{Preliminary}

Consider a bounded function
\[ \phi : \mathbb{R} \rightarrow \mathbb{C}. \]

The function \( \phi^j \) is defined by dilation of \( \phi \):
\[ \phi^j(x) = \sqrt{2^j} \phi(2^j x) \text{ for } j \in \mathbb{Z} \text{ and } j \leq 0. \]

It is a function that becomes increasingly stretched as \( j \) decreases. The functions \( \phi^j_k \in l^2 \) are defined by
\[ \phi^j_k(n) = \phi^j(n - k2^{-j}) \text{ with } n, k \in \mathbb{Z} \text{ and } j \leq 0, \]
\[ = \sqrt{2^j} \phi(2^j n - k). \]

Hence, the functions \( \phi^j_k \) are constructed by firstly shifting \( \phi^j \) over a discrete distance \( k2^{-j} \), which increases as \( j \) decreases, and finally by sampling the result at discrete sampling points \( n \).

The vectorspace \( M_j \subset l^2 \) will be defined as
\[ M_j = \text{span}\{\phi^j_k\}_{k=-\infty,\infty}. \]

A function \( \phi \) will be called a scaling function if
\[ M_{j-1} \subset M_j \quad \forall j \leq 0 \tag{3.3.1} \]

and if \( \{\phi^j_k\}_{k=-\infty,\infty} \) constitutes an orthonormal basis of \( M_j \):
\[ (\phi^j_k, \phi^j_m) = \begin{cases} 1 & \text{if } m = k, \\ 0 & \text{if } m \neq k. \end{cases} \tag{3.3.2} \]
Orthonormality of \( \{ \phi_k^0 \}_{k=-\infty, \infty} \) ensures that
\[
M_0 = l^2 .
\] (3.3.3)

The projection of \( f \in l^2 \) onto \( M_j \),
\[
P_{M_j}(f) = \sum_{k=-\infty}^{\infty} (f, \phi_k^j) \phi_k^j ,
\]
will be called the representation of \( f \) at scale \( j \). \( (f, \phi_k^j) \) will be called a multiresolution coefficient. The fact that we are actually dealing with a multiresolution representation stems from eq.(3.3.1) and the way in which the functions \( \phi_k^j \) are constructed. Note the similarity between the right-hand side of eq.(3.2.3) and the multiresolution coefficients. Eq.(3.2.1) will appear after the next three theorems have been established.

**Theorem 1** Suppose that \( \phi \in BL^2(\mathbb{R}) \) is a left or right continuous function for which \( \phi \) is continuous in 0, with \( \phi(0) = 1 \), and suppose that, for certain \( T > 0 \) and \( M > 0 \),
\[
|\phi(x)| \leq \frac{M}{|x|^\alpha} \quad \forall \quad |x| \geq T \text{ with } \alpha > \frac{2}{3} .
\] (3.3.4)

If such a function \( \phi \) is a scaling function then it satisfies
\[
\phi\left(\frac{x}{2}\right) = 2 \sum_{n=-\infty}^{\infty} h(n)\phi(x - n) ,
\] (3.3.5)

where \( h : \mathbb{Z} \to \mathbb{C} \) is a function with discrete-time Fourier transform \( H \), such that

\( H \) is continuous in 0 ,
\[
H(0) = 1 ,
\] (3.3.6)

\[
H(\omega)e^{i\theta(\omega)} + H(\omega + \pi)e^{i\theta(\omega + \pi)} = e^{i\theta(2\omega)} ,
\]
where \( e^{i\theta(\omega)} = e^{i\theta(\omega + 2\pi)} \),
\[
|H(\omega)|^2 + |H(\omega + \pi)|^2 = 1 .
\] (3.3.7)
If \( \alpha > 1 \) the condition on the continuity of \( \tilde{\phi} \) in 0 is satisfied automatically, because in that case the scaling function \( \phi \in L^1(\mathbb{R}) \), which implies that \( \tilde{\phi} \) is a continuous function (Weaver[36]).

**Theorem 2** A left or right continuous function \( \phi \in BL^2(\mathbb{R}) \) is a scaling function if it satisfies eq.(3.3.5), where the discrete-time Fourier transform \( H \) of \( h \) satisfies eq.(3.3.8), and if in addition

\[
\sum_{n=-\infty}^{\infty} \phi(n-k)\phi^*(n) = \begin{cases} 1 & k = 0, \\ 0 & k \neq 0. \end{cases} \tag{3.3.9}
\]

Eq.(3.3.9) ensures the orthonormality of the \( \phi_k^0 \)'s. The next theorem states that if \( \tilde{\phi} \in L^1(\mathbb{R}) \), this property can be obtained directly from properties of \( H \).

**Theorem 3** A function \( \phi \in BL^2(\mathbb{R}) \) is a scaling function if \( \tilde{\phi} \in L^1(\mathbb{R}) \) and if it satisfies eq.(3.3.5), where the discrete-time Fourier transform \( H \) of \( h \) has to be differentiable in 0 and satisfies eqs.(3.3.7), (3.3.8) and

\[
H(\omega) \neq 0 \quad \forall \omega \in \left[ -\frac{\pi}{2}, \frac{\pi}{2} \right]. \tag{3.3.10}
\]

We note that \( \tilde{\phi} \in L^1(\mathbb{R}) \) implies that \( \phi \) is continuous (Weaver[36]). The proof of Theorems 1,2 and 3 is given in Appendix 3.A.

Fourier transformation of eq.(3.3.5) yields

\[
\tilde{\phi}(2\omega) = H(\omega)\tilde{\phi}(\omega). \tag{3.3.11}
\]

If eq.(3.3.6) is satisfied the previous equation can be written as

\[
\tilde{\phi}(\omega) = \prod_{k=1}^{\infty} H(2^{-k}\omega). \tag{3.3.12}
\]

### 3.3.2 Construction

Theorem 1 describes properties of a scaling function. Theorem 2 and Theorem 3 describe how a scaling function can be constructed from a discrete filter \( h \). The phase factor \( e^{i\theta(\omega)} \) in eq.(3.3.7) can be eliminated by constructing \( H \) such that \( \theta(\omega) = 0 \). If this is the case eq.(3.3.7) is equivalent with
\[ h(2n) = \begin{cases} \frac{1}{2} & \text{if } n = 0, \\ 0 & \text{if } n \neq 0. \end{cases} \]

As is shown in Appendix 3.A, the relation between \( h \) and the \( \phi^j_k \)'s is expressed by

\[ h(k - 2m) = \frac{1}{\sqrt{2}} (\phi^j_m^{-1}, \phi^j_k), \]

with \( j \) arbitrary but \( j \leq 0 \).

From eq.(3.3.5) we obtain, straightforwardly,

\[
\phi^j_m^{-1}(n) = \sum_{k=-\infty}^{\infty} (\phi^j_m^{-1}, \phi^j_k) \phi^j_k(n)
\]

\[ = \sqrt{2} \sum_{k=-\infty}^{\infty} h(k - 2m) \phi^j_k(n) \quad (3.3.13) \]

and

\[
(f, \phi^j_m^{-1}) = \sqrt{2} \sum_{k=-\infty}^{\infty} h(k - 2m) (f, \phi^j_k). \quad (3.3.14)
\]

Eq.(3.3.14) is similar to eq.(3.2.1) and shows that the multiresolution coefficients at scale \( j - 1 \) can be obtained by smoothing and downsampling those at scale \( j \). Note that \( h \) is a smoothing or low-pass filter because of eqs.(3.3.6) and (3.3.8).

Eq.(3.3.14) is the answer to the objectives formulated in the introduction for a multiresolution representation. It gives the discrete filter that connects two scales and it gives the discrete filter that connects a certain scale to the original signal. The starting coefficients in eq.(3.3.14) are

\[ (f, \phi^j_k) \text{ for } k = -\infty, \infty. \]

Under the conditions in Theorems 2 or 3, the \( \phi^j_0 \)'s are given by

\[ \phi^j_0(n) = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{i\theta} e^{in} e^{i\omega(n-k)} \, d\omega. \]

For the construction of a multiresolution representation in which only the coefficients are required it is not necessary to know or calculate \( \phi \), provided the orthonormality of the \( \phi^j_0 \)'s is guaranteed and (left or right) continuity of \( \phi \).
can be verified. If $H$ satisfies the conditions of Theorem 3 only the absolute integrability of $\tilde{\phi}$ remains to be verified. This subject will return in Section 3.7.

**Example**

$$H(\omega) = \chi_{[-\pi/2, \pi/2]}(\omega), \; 2\pi\text{-periodically continued.} \quad (3.3.15)$$

$H$ satisfies the conditions of Theorem 3. With eq.(3.3.12), it follows that

$$\tilde{\phi}(\omega) = \chi_{[-\pi, \pi]}(\omega)$$

and inverse Fourier transformation yields

$$\phi(x) = \frac{\sin(\pi x)}{\pi x}.$$  

Since $\tilde{\phi} \in L^1(\mathbb{R})$, all conditions in Theorem 3 are met and we conclude that $\phi$ is a scaling function.

### 3.4 Wavelet representation for $l^2$

#### 3.4.1 Preliminary

Suppose that we have a bounded function

$$\psi : \mathbb{R} \rightarrow \mathcal{C}.$$  

We define

$$\psi^j_k(n) = \sqrt{2^j} \psi(2^j n - k), \text{ with } n, k, j \in \mathbb{Z} \text{ and } j \leq 0.$$  

The vectorspace $W_j$ will be defined as

$$W_j = \text{span}\{\psi^j_k\}_{k=-\infty, \infty}.$$  

A function $\psi$ will be called a wavelet function if

$$M_{j+1} = M_j \oplus W_j \quad (3.4.1)$$

and if $\{\psi^j_k\}_{k=-\infty, \infty}$ constitutes an orthonormal basis of $W_j$,

$$(\psi^j_k, \psi^j_m) = \begin{cases} 1 & \text{if } m = k, \\ 0 & \text{if } m \neq k. \end{cases}$$
In this case the $\psi_j^i$'s will be called wavelets.

$M_j$ consists of smoother functions than $M_{j+1}$, as follows from eq.(3.3.13). The vectorspace $W_j$ contains the information complementary to $M_j$ to build $M_{j+1}$. Hence, $W_j$ represents a band of frequencies that connects $M_j$ and $M_{j+1}$.

The projection of a function $f \in l^2$ onto $W_j$,

$$PW_j(f) = \sum_{k=-\infty}^{\infty} (f, \psi_k^j) \psi_k^j,$$

will be referred to as the wavelet representation of $f$ in subband $j$. $(f, \psi_k^j)$ is called a wavelet coefficient. $(f, \psi_k^j)$ is similar to eq.(3.2.4). From eqs.(3.3.1), (3.3.3) and (3.4.1) it follows that

$$l^2 = \bigcup_{j=-1}^{\infty} W_j$$

and that \(\{\psi_j^i\}_{k=-\infty}^{\infty} \) is an orthonormal basis of $l^2$.

The relation between $\phi$ and $\psi$ is expressed in the next theorem.

**Theorem 4** If $\phi$ is a left continuous scaling function then a left continuous $\psi \in BL^2(\mathbb{R})$ is a wavelet function if

$$\psi\left(\frac{x}{2}\right) = 2 \sum_{n=-\infty}^{\infty} g(n) \phi(x - n), \quad (3.4.2)$$

where $g : \mathbb{Z} \to \mathbb{C}$ is a function with discrete-time Fourier transform $G$, that satisfies

$$|G(\omega)|^2 + |H(\omega)|^2 = 1 \quad (3.4.3)$$

and

$$G(\omega)H^*(\omega) + G(\omega + \pi)H^*(\omega + \pi) = 0 \quad (3.4.4)$$

If $\phi$ is a left continuous scaling function, which satisfies eq.(3.3.4), then a left continuous $\psi \in BL^2(\mathbb{R})$, which also satisfies eq.(3.3.4), is a wavelet function only if eq.(3.4.2), eq.(3.4.3) and eq.(3.4.4) are satisfied.
This theorem is also valid for right continuous $\phi$ and $\psi$. It is proved in Appendix 3.B.

Eq. (3.4.3) shows that $H$ and $G$ are QMF-filters. From eqs. (3.3.8), (3.4.3) and (3.4.4), it is straightforward to derive that

$$G(\omega) = H^*(\omega + \pi) e^{i\pi(\omega)},$$

where $e^{i\pi(\omega + \pi)} = e^{-i\pi(\omega + \pi)}$. \hfill (3.4.6)

$G$ also satisfies

$$|G(\omega)|^2 + |G(\omega + \pi)|^2 = 1$$

and from eq. (3.4.2) we obtain

$$\tilde{\psi}(2\omega) = G(\omega)\tilde{\phi}(\omega).$$ \hfill (3.4.7)

3.4.2 Construction

As is shown in Appendix 3.B the discrete filter $g$ satisfies

$$g(k - 2m) = \frac{1}{\sqrt{2}} (\psi_m^{j-1}, \phi_k^j),$$

with $j$ arbitrary but $j \leq 0$.

From eq. (3.4.2) it follows that

$$\psi_m^{j-1}(n) = \sum_{k=-\infty}^{\infty} (\psi_m^{j-1}, \phi_k^j) \phi_k^j(n)$$

$$= \sqrt{2} \sum_{k=-\infty}^{\infty} g(k - 2m) \phi_k^j(n)$$

and

$$(f, \psi_m^{j-1}) = \sqrt{2} \sum_{k=-\infty}^{\infty} g(k - 2m) (f, \phi_k^j).$$ \hfill (3.4.9)

Eq. (3.4.9) is similar to eq. (3.2.2) and shows that the wavelet coefficients in subband $j - 1$ can be obtained by high-pass filtering and downsampling the multiresolution coefficients at scale $j$. Eq. (3.3.14) and eq. (3.4.9) provide an algorithm for the calculation of the wavelet coefficients for all $j$ (fig. 3.4.1). The wavelet coefficients in subband $j - 1$ can be obtained from the multiresolution coefficients at scale $j$ with eq. (3.4.9). With eq. (3.3.14) the coefficients at scale $j - 1$ can be calculated as input for further calculation.
3.4.3 Reconstruction

Eq.(3.4.1) means that the multiresolution coefficients at scale $j$ can be obtained from the multiresolution coefficients at scale $j - 1$ and the wavelet coefficients in subband $j - 1$. In Appendix 3.B it is proved that

$$\phi^j_m(n) = \sum_{k=-\infty}^{\infty} (\phi^i_m, \phi^i_{k-1}) \phi^{i-1}_k(n) + (\phi^i_m, \psi^i_{k-1}) \psi^{i-1}_k(n)$$

$$= \sqrt{2} \sum_{k=-\infty}^{\infty} h^*(m - 2k) \phi^{i-1}_k(n) + g^*(m - 2k) \psi^{i-1}_k(n)$$

and consequently

$$(f, \phi^i_m) = \sum_{k=-\infty}^{\infty} (\phi^i_m, \phi^i_{k-1}) (f, \phi^i_{k-1}) + (\phi^i_m, \psi^i_{k-1}) (f, \psi^i_{k-1})$$

$$= \sqrt{2} \sum_{k=-\infty}^{\infty} h^*(m - 2k) (f, \phi^i_{k-1}) + g^*(m - 2k) (f, \psi^i_{k-1}) . \ (3.4.10)$$

Eq.(3.4.10) provides the algorithm for the reconstruction of the original signal from its wavelet coefficients (fig.3.4.2).
Example

For the filter $H$ in eq.(3.3.15) and the corresponding scaling function $\phi$ in the example in the previous section, a possible $G$ is given by

$$G(\omega) = H^*(\omega + \pi)e^{-i\omega}$$

$$= \left(\chi[-\pi,\pi/2] + \chi(\pi/2,\pi]\right) e^{-i\omega}, \text{2\pi-periodically continued.}$$

In this case we have

$$\bar{\psi}(\omega) = \left[\chi[-\pi/2,\pi/2]\left(\omega\right) - \chi[-\pi/4,\pi/4]\left(\omega\right)\right] e^{-i\frac{\omega}{2}}.$$ 

The wavelet function is

$$\psi(x) = \frac{\sin\left(\frac{1}{2}\pi(x - \frac{1}{2})\right)}{\pi(x - \frac{1}{2})} - \frac{\sin\left(\frac{1}{4}\pi(x - \frac{1}{2})\right)}{\pi(x - \frac{1}{2})}.$$
3.5 General formula for the QMF-filters

From eqs. (3.3.7), (3.3.8), (3.4.5) and (3.4.6) a general formula for both $H$ and $G$ can be derived.

**Theorem 5** The general formula for $H$ is

$$H(\omega) = \frac{1}{2}(1 + e^{i\zeta(\omega)})e^{i[\theta(2\omega) - \theta(\omega)]},$$

where $e^{i\theta(\omega)} = e^{i\theta(\omega + 2\pi)} \quad (3.5.1)$

and where $\zeta$ should be such that

$$e^{i\zeta(\omega + 2\pi)} = e^{i\zeta(\omega)}$$

and

$$e^{i\zeta(\omega + \pi)} = -e^{i\zeta(\omega)} \quad (3.5.2)$$

The general formula for $G$ is

$$G(\omega) = \frac{1}{2}(1 - e^{-i\zeta(\omega)})e^{-i[\theta(2\omega) - \theta(\omega + \pi)]}e^{i\eta(\omega)},$$

where $e^{i\eta(\omega)} = -e^{i\eta(\omega + \pi)}.$

These general formulas are convenient in the design of QMF-filters and study of their properties as we will see in the following sections. The proof of Theorem 5 will be given after the next example.

**Example**

For the filter $H$ in eq. (3.3.15) we find:

$$\theta(\omega) = 0$$

and

$$\zeta(\omega) = \begin{cases} 
0 & \text{if } 0 \leq |\omega| \leq \frac{\pi}{2}, \\
\pi & \text{if } \frac{\pi}{2} < |\omega| \leq \pi, \\
2\pi & \text{periodic}.
\end{cases}$$
Proof of Theorem 5

Define

\[ V(\omega) = e^{-i[\theta(2\omega)-\theta(\omega)]} H(\omega) = \sum_{n=-\infty}^{\infty} v(n) e^{-i\omega n} . \]

Then,

\[ V(\omega + \pi) = e^{-i[\theta(2\omega)-\theta(\omega+\pi)]} H(\omega + \pi) = \sum_{n=-\infty}^{\infty} v(n) (-1)^n e^{-i\omega n} . \]

For \( V \), eq.(3.3.7) becomes

\[ V(\omega) + V(\omega + \pi) = 1 \]

and this implies that

\[ \sum_{n=-\infty}^{\infty} v(2n) e^{-i\omega 2n} = \frac{1}{2} . \quad (3.5.3) \]

This leads to the conclusion that

\[ v(2n) = \begin{cases} \frac{1}{2} & \text{if } n = 0 , \\ 0 & \text{if } n \neq 0 . \end{cases} \quad (3.5.4) \]

If eq.(3.3.7), and consequently eq.(3.5.4), is satisfied eq.(3.3.8) becomes

\[ |v(0) + \sum_{n=-\infty}^{\infty} v(2n + 1) e^{-i\omega(2n+1)}|^2 + \]

\[ |v(0) - \sum_{n=-\infty}^{\infty} v(2n + 1) e^{-i\omega(2n+1)}|^2 = 1 . \]

This equation results in

\[ \left| \sum_{n=-\infty}^{\infty} v(2n + 1) e^{-i\omega(2n+1)} \right|^2 = \frac{1 - 2|v(0)|^2}{2} \]

\[ = \frac{1}{4} . \]

Hence

\[ \sum_{n=-\infty}^{\infty} v(2n + 1) e^{-i\omega(2n+1)} = \frac{1}{2} e^{i\zeta(\omega)} , \quad (3.5.5) \]
where \( \zeta(\omega) \) satisfies
\[
e^{i\zeta(\omega+2\pi)} = e^{i\zeta(\omega)}
\]
and
\[
e^{i\zeta(\omega+\pi)} = -e^{i\zeta(\omega)}.
\]

From eqs. (3.5.3) and (3.5.5) we obtain for \( V \):
\[
V(\omega) = \frac{1}{2} + \frac{1}{2} e^{i\zeta(\omega)}.
\]

Consequently, the general formula for \( H \) is given by
\[
H(\omega) = \frac{1}{2} (1 + e^{i\zeta(\omega)}) e^{i[\theta(2\omega) - \theta(\omega)]},
\]

where \( \zeta \) satisfies the conditions in eq. (3.5.2) and where \( e^{i\theta(\omega)} = e^{i\theta(\omega+2\pi)} \).

The general formula for \( G \) now simply follows from eqs. (3.4.5) and (3.4.6). \( \square \)

### 3.6 Finite duration QMF-filters

In practice it is only possible to process signals of finite duration. It is therefore important to determine QMF-filters, scaling functions and wavelet functions that have finite duration.

**Theorem 6** Except for a possible integer time shift, the only QMF-filters, scaling function and wavelet function that have finite duration are
\[
\begin{align*}
\hat{h}(n) &= \begin{cases} 
\frac{1}{2} & \text{if } n = 0, \\
\frac{1}{2} & \text{if } n = 1, \\
0 & \text{elsewhere,}
\end{cases} \\
g(n) &= \begin{cases} 
-\frac{1}{2} & \text{if } n = 0, \\
\frac{1}{2} & \text{if } n = 1, \\
0 & \text{elsewhere,}
\end{cases} \\
\phi(x) &= \begin{cases} 
1 & \text{for } 0 \leq x < 1, \\
0 & \text{for } x < 0 \text{ or } x \geq 1,
\end{cases} \\
\psi(x) &= \begin{cases} 
-1 & \text{for } 0 \leq x < \frac{1}{2}, \\
1 & \text{for } \frac{1}{2} \leq x < 1, \\
0 & \text{for } x < 0 \text{ or } x \geq 1.
\end{cases}
\end{align*}
\]

The corresponding orthonormal wavelet basis of \( l^2 \) is usually called the Haar basis.
Proof of Theorem 6
From eqs. (3.5.1) and (3.5.2), we obtain that if $h$ is to be of finite duration,
\[ e^{i\xi(\omega)} = e^{-i(2k+1)\omega} , \text{ with } k \text{ an arbitrary integer,} \]
and
\[ e^{i[\theta(2\omega)-\theta(\omega)]} = e^{-im\omega} , \text{ with } m \text{ an arbitrary integer.} \]

In this way, $H$ becomes
\[ H(\omega) = \frac{1}{2}(1 + e^{-i(2k+1)\omega})e^{-im\omega} . \] (3.6.1)

The factor $e^{-im\omega}$ represents a mere integer time shift and can be omitted without loss of generality. The corresponding $h$ is given by
\[ h(n) = \begin{cases} 
\frac{1}{2} & \text{if } n = 0 , \\
\frac{1}{2} & \text{if } n = 2k + 1 , \\
0 & \text{elsewhere,} 
\end{cases} \] (3.6.2)

where $k$ is an arbitrary integer. The discrete-time Fourier transform of $h$ can also be written as
\[ H(\omega) = e^{-i(k+\frac{1}{2})\omega} \cos((k + \frac{1}{2}) \omega) . \]

The unique $\tilde{\phi}$ that corresponds to $H$ results from eq. (3.3.11) or eq. (3.3.12) and is given by
\[ \tilde{\phi}(\omega) = e^{-i(k+\frac{1}{2})\omega} \frac{\sin((k + \frac{1}{2}) \omega)}{(k + \frac{1}{2}) \omega} . \]

Inverse Fourier transformation reveals that the corresponding left continuous $\phi$ is given by
\[ \phi(x) = \begin{cases} 
1 & \text{for } 0 \leq x < 2k + 1 , \\
0 & \text{for } x < 0 \text{ or } x \geq 2k + 1 . 
\end{cases} \] (3.6.3)

With reference to Theorem 2, the orthonormality of the $\phi^0_k$'s, expressed in eq. (3.3.9), remains to be verified. It follows that $k = 0$ in eqs. (3.6.1), (3.6.2), and (3.6.3). Hence,
\[ H(\omega) = \frac{1}{2}(1 + e^{-i\omega}) , \] (3.6.4)
\[ h(n) = \begin{cases} 
\frac{1}{2} & \text{if } n = 0, \\
\frac{1}{2} & \text{if } n = 1, \\
0 & \text{elsewhere.} 
\end{cases} \]

and

\[ \phi(x) = \begin{cases} 
1 & \text{for } 0 \leq x < 1, \\
0 & \text{for } x < 0 \text{ or } x \geq 1. 
\end{cases} \tag{3.6.5} \]

Note that

\[ \phi(n) = \begin{cases} 
1 & \text{for } n = 0, \\
0 & \text{for } n \neq 0. 
\end{cases} \]

Substitution of eq. (3.6.4) in eq. (3.4.5) results in

\[ G(\omega) = \frac{1}{2} (1 - e^{i\omega}) e^{i\eta(\omega)}. \]

\( g \) has a finite impulse response only if \( \eta(\omega) = p\omega \), where \( p \) is an (odd) integer. The factor \( e^{i\eta(\omega)} \) only represents an integer time shift. \( \eta(\omega) \) taken to be \( -\omega \). This leads to

\[ G(\omega) = \frac{1}{2} (-1 + e^{-i\omega}). \]

\( g \) is given by

\[ g(n) = \begin{cases} 
-\frac{1}{2} & \text{if } n = 0, \\
\frac{1}{2} & \text{if } n = 1. 
\end{cases} \]

With eq. (3.4.2), eq. (3.6.5) implies that

\[ \psi(x) = \begin{cases} 
-1 & \text{for } 0 \leq x < \frac{1}{2}, \\
1 & \text{for } \frac{1}{2} \leq x < 1, \\
0 & \text{for } x < 0 \text{ or } x \geq 1. 
\end{cases} \]
3.7 Continuity of the scaling and wavelet function

3.7.1 Decay rate

In Theorem 2 left and/or right continuity of the scaling function $\phi$ is presumed. In Theorem 3 its Fourier transform $\tilde{\phi}$ is presumed to be absolutely integrable. Both properties can be ensured by a sufficiently large decay rate of $\tilde{\phi}$ at infinity. This leads to additional conditions for $H$ (Mallat[23]).

Suppose that $H \in C^q$. If

$$\frac{d^n H}{d\omega^n}(0) = 0 \text{ for } 1 \leq n \leq q$$  \hspace{1cm} (3.7.1)

then, by evaluation of the $n$-th derivative of eq.(3.3.8) for every $0 \leq n \leq q$ in $\omega = 0$, we arrive at

$$\frac{d^n H}{d\omega^n}(\pi) = 0 \text{ for } 0 \leq n \leq q - 1 .$$

Hence, $H$ can be written as

$$H(\omega) = (\cos(\frac{\omega}{2}))^q T(\omega) ,$$

in which

$$|T(\omega)| \leq A < \infty .$$  \hspace{1cm} (3.7.2)

Daubechies[10] has shown that

$$\prod_{k=1}^{\infty} |T(2^{-k}\omega)| = O(|\omega|^{|\log_2(A)|}) \text{ as } |\omega| \to \infty .$$

Since

$$\prod_{k=1}^{\infty} \cos(2^{-k}\frac{\omega}{2}) = \frac{\sin(\frac{1}{2}\omega)}{\frac{1}{2}\omega} ,$$

this means that

$$|\tilde{\phi}(\omega)| = O(|\omega|^{-q+|\log_2(A)|}) \text{ as } |\omega| \to \infty .$$  \hspace{1cm} (3.7.3)

The decay of $\tilde{\phi}$ is determined by eqs.(3.7.1), (3.7.2) and (3.7.3). The decay rate of $\tilde{\psi}$ is the same as for $\tilde{\phi}$, as follows from eq.(3.4.8). A proper decay rate of $\tilde{\phi}$ and $\tilde{\psi}$ can guarantee continuity and even differentiability.

Suppose that $|\tilde{\phi}(\omega)| = O(\frac{1}{|\omega|^{\alpha}}) \text{ as } |\omega| \to \infty .$
- If $\alpha > 1$ then $\tilde{\phi}$ is absolutely integrable, which implies that $\phi$ is (uniformly) continuous (Weaver[36]).

- If $\alpha > n + \frac{1}{2}$ then $\omega^n \tilde{\phi}(\omega)$ is square integrable, which implies that $\frac{d^n \phi}{dx^n}$ is square integrable.

- If $\alpha > n + 1$ then $\frac{d^n \phi}{dx^n}$ is (uniformly) continuous.

An estimate for the decay rate of $h$ can be obtained as well. As is shown in Appendix 3.C, $H \in C^q$ implies that

$$|h(n)| = o\left(\frac{1}{|n|^{q+1}}\right) \text{ as } |n| \to \infty.$$  

### 3.7.2 Consequence

The results of the previous subsection have to be transferred to conditions for the function $\zeta$ in the general formula in eq.(3.5.1). Recall eqs.(3.5.1) and (3.5.2), with $\theta(\omega) = 0$,

$$H(\omega) = \frac{1}{2}(1 + e^{i\zeta(\omega)}),$$

where

$$e^{i\zeta(\omega + 2\pi)} = e^{i\zeta(\omega)}$$

and

$$e^{i\zeta(\omega + \pi)} = -e^{i\zeta(\omega)}.$$

A number of observations can be made:

- $H \in C^q \iff \zeta \in C^q$.

- $\zeta(0) = 0 \Rightarrow H(0) = 1$.

- $H(\omega) \neq 0$ for $\omega \in [-\frac{\pi}{2}, \frac{\pi}{2}]$ \iff $\zeta(\omega) \neq (2k + 1)\pi$ for $\omega \in [-\frac{\pi}{2}, \frac{\pi}{2}]$.

- $\frac{d^n H}{d\omega^n}(0) = 0$ for $1 \leq n \leq q \Leftrightarrow \frac{d^n \zeta}{d\omega^n}(0) = 0$ for $1 \leq n \leq q$. 
\( \zeta \) satisfies eq.(3.5.2) if
\[
\zeta(\omega) = \zeta(\omega - k\pi) + k\pi ,
\]
with \( -\frac{\pi}{2} + k\pi < \omega \leq \frac{\pi}{2} + k\pi . \)  \hspace{1cm} (3.7.4)

Eq.(3.7.4) evaluated at \( \omega = -\frac{\pi}{2} \) gives
\[
\zeta(-\frac{\pi}{2}) = \zeta(\frac{\pi}{2}) - \pi \hspace{1cm} (3.7.5)
\]
The behavior of \( \zeta \) and its derivatives at the boundary points \( \frac{\pi}{2} + k\pi \) has to be examined carefully with respect to continuity (\( \zeta \in C^q \)). Eq.(3.7.4) implies that, for \( 1 \leq n \leq q \),
\[
\frac{d^n\zeta}{d\omega^n}(\omega) = \frac{d^n\zeta}{d\omega^n}(\omega - k\pi)
\]
for \( -\frac{\pi}{2} + k\pi < \omega \leq \frac{\pi}{2} + k\pi . \)  \hspace{1cm} (3.7.6)

Eq.(3.7.6) evaluated at \( \omega = -\frac{\pi}{2} \) gives
\[
\frac{d^n\zeta}{d\omega^n}(-\frac{\pi}{2}) = \frac{d^n\zeta}{d\omega^n}(\frac{\pi}{2}) \text{ for } 1 \leq n \leq q . \hspace{1cm} (3.7.7)
\]
The discrete-time filter \( h \) is real valued if \( H(\omega) = H^*(-\omega) \), which is the case if \( \zeta \) is an odd function. This has some interesting consequences. If \( \zeta \) is odd this leads to
\[
\frac{d^n\zeta}{d\omega^n}(\omega) = (-1)^{n+1} \frac{d^n\zeta}{d\omega^n}(-\omega) \text{ for } n \geq 0 . \hspace{1cm} (3.7.8)
\]
From eq.(3.7.8) and eq.(3.7.5) it follows that
\[
\zeta(\frac{\pi}{2}) = \frac{\pi}{2}
\]
and eqs.(3.7.8) and (3.7.7) result in
\[
\frac{d^n\zeta}{d\omega^n}(\frac{\pi}{2}) = (-1)^{n+1} \frac{d^n\zeta}{d\omega^n}(\frac{\pi}{2}) \text{ for } 1 \leq n \leq q . \hspace{1cm} (3.7.9)
\]
Obviously,
\[
\frac{d^n\zeta}{d\omega^n}(\frac{\pi}{2}) = 0 \text{ for } n \text{ even.}
\]
Note that for \( n \) odd eq.(3.7.7) is satisfied automatically.
3.8 Construction of an example

The equations that $H$ or $\zeta$ has to satisfy are not really complicated. However, we have not been able to find filters $H$, that differ essentially from the examples we have already given, for which $h$, $\phi$ or $\dot{\phi}$ can be calculated analytically. Instead, we will construct a QMF-filter $H$ for which $h$, $\phi$ and $\dot{\phi}$ will be calculated numerically. From $H$ we obtain $G$ for which $g$, $\psi$ and $\dot{\psi}$ will be calculated numerically as well. $H$ and $G$ are given by

$$H(\omega) = \frac{1}{2} (1 + e^{i\zeta(\omega)}) ,$$
$$G(\omega) = \frac{1}{2} (1 - e^{-i\zeta(\omega)}) e^{-i\omega} .$$

In this example we will use a function $\zeta$ which satisfies

$$\zeta(\omega) = \frac{\pi}{2} \left( \frac{2}{\pi} \omega \right)^6 \left( 3 \left( \frac{2}{\pi} \omega \right)^2 - 9 \left( \frac{2}{\pi} \omega \right) + 7 \right) \text{ for } \omega \in [0, \frac{\pi}{2}] ,$$

$$\zeta(\omega) = -\zeta(-\omega) \text{ for } \omega \in \left[ -\frac{\pi}{2}, 0 \right) \text{ and}$$

$$\zeta \text{ is given by eq.}(3.7.4) \text{ for } \omega \text{ outside } \left[ -\frac{\pi}{2}, \frac{\pi}{2} \right] .$$

It is straightforward to verify that $\zeta \in C^5$ and that

$$\frac{d^n\zeta}{d\omega^n}(0) = 0 \text{ for } 0 \leq n \leq 5 ,$$

$$\frac{d^n\zeta}{d\omega^n}(\frac{\pi}{2}) = 0 \text{ for } n = 2, 4 ,$$

$$\frac{d^n\zeta}{d\omega^n}(\pi) = 0 \text{ for } 1 \leq n \leq 5 .$$

Because $H \in C^5$ and $G \in C^5$, the discrete-time filters $h$ and $g$ decay as $\frac{1}{|n|^6}$.

The amplitude of $H$ and $G$ is depicted in fig.3.8.1. Their real and imaginary parts follow in fig.3.8.2 and fig.3.8.3. Because

$$H(\omega) = H^*(-\omega) \text{ and } G(\omega) = G^*(-\omega) , \quad (3.8.1)$$

their inverse Fourier transforms $h$ and $g$ are real valued. They have been calculated using a FFT-routine. $h$ and $g$ are given in fig.3.8.4. It can be
observed that both \( h \) and \( g \) decay rapidly, as was to be expected, and that they are non-symmetric.

The decay rate of \( \tilde{\phi} \) and \( \tilde{\psi} \) can be estimated with eq.(3.7.3). Numerically, we have found that \( A \leq 5.13 \) in eq.(3.7.2). Hence,

\[ |\tilde{\phi}| = O\left(\frac{1}{|\omega|^{2.64}}\right) \quad \text{as} \quad |\omega| \to \infty \]

and

\[ |\tilde{\psi}| = O\left(\frac{1}{|\omega|^{2.64}}\right) \quad \text{as} \quad |\omega| \to \infty . \]

This implies that both \( \phi \) and \( \psi \) have a continuous first derivative and a square integrable second derivative. \( \tilde{\phi} \) has been calculated by numerical evaluation of the infinite product in eq.(3.3.12). \( \tilde{\psi} \) is obtained using eq.(3.4.8). The amplitude of \( \tilde{\phi} \) and \( \tilde{\psi} \) are depicted in fig.3.8.5. Their real and imaginary part are displayed in fig.3.8.6 and fig.3.8.7. We observe that the width of \( \tilde{\psi} \) is approximately twice that of \( \tilde{\phi} \), which is a consequence of eq.(3.4.8). From eq.(3.8.1) it follows that

\[ \tilde{\phi}(\omega) = -\tilde{\phi}(-\omega) \quad \text{and} \quad \tilde{\psi}(\omega) = -\tilde{\psi}(-\omega) . \]

Consequently, the scaling function \( \phi \) and the wavelet function \( \psi \) are real valued. They have been calculated from \( \tilde{\phi} \) and \( \tilde{\psi} \) using a FFT-routine and have been depicted in fig.3.8.8. It can be observed that the width of \( \phi \) is approximately twice that of \( \psi \).

### 3.9 Conclusions

A theory for orthonormal wavelet bases for discrete-time signals has been developed. These wavelets are obtained by dilating, shifting and sampling a basic function, called the wavelet function. A wavelet decomposition can be performed using an exact reconstruction subband filterbank of which the QMF-filters uniquely determine the wavelet function. The QMF-filters have to satisfy only one condition additional to the conditions they have to satisfy in the theory for \( L^2(\mathbb{R}) \). A general formula for the QMF-filters exists that can be used for their design. The Haar-basis appears to be the only wavelet basis that has finite duration QMF-filters. The smoothness of the wavelet function can be controlled by the smoothness of the QMF-filters; differentiability conditions on the QMF-filters result in differentiability of the wavelet function.
Figure 3.8.1: Amplitude of the QMF-filters $H$ (left) and $G$ (right) in the frequency domain.

Figure 3.8.2: Real part of the QMF-filters $H$ (left) and $G$ (right) in the frequency domain.
Figure 3.8.3: Imaginary part of the QMF-filter $H$ (left) and $G$ (right) in the frequency domain.

Figure 3.8.4: The QMF-filters $h$ (left) and $g$ (right) in the time domain.
Figure 3.8.5: Amplitude of the scaling function (left) and the wavelet function (right) in the frequency domain.

Figure 3.8.6: Real part of the scaling function (left) and the wavelet function (right) in the frequency domain.
Figure 3.8.7: Imaginary part of the scaling function (left) and the wavelet function (right) in the frequency domain.

Figure 3.8.8: The scaling function $\phi$ (left) and the wavelet function $\psi$ in the time domain.
3.A Proof of Theorem 1, 2 and 3

3.A.1 Proof of Theorem 1

It will be proved that if $\phi$ is a left or right continuous scaling function, it satisfies eq.(3.3.5) and subsequently eqs.(3.3.6), (3.3.7) and (3.3.8) will be proved.

Eq.(3.3.5)

Since $\phi$ is a scaling function it follows that

$$\phi_{m}^{-1}(n) = \sum_{k=-\infty}^{\infty} (\phi_{m}^{j-1}, \phi_{k}^{j}) \phi_{k}^{j}(n).$$  \hspace{1cm} (3.A.1)

We define

$$h_{j}(k) = \frac{1}{2} \sum_{n=-\infty}^{\infty} 2^{j} \phi(2^{j-1}n) \phi^{*}(2^{j}n - k).$$  \hspace{1cm} (3.A.2)

The right-hand-side of this expression is a Riemann sum and therefore we know that

$$\lim_{j \to -\infty} h_{j}(k) = \frac{1}{2} \int_{-\infty}^{\infty} \phi(\frac{1}{2} x) \phi^{*}(x - k) \, dx,$$

$$= h(k).$$ \hspace{1cm} (3.A.3)

With $m = 0$, eq.(3.A.1) can be written as

$$\phi(\frac{1}{2} 2^{j}n) = 2 \sum_{k=-\infty}^{\infty} h_{j}(k) \phi(2^{j}n - k).$$

Consider,

$$x_{j} = n_{j}2^{j} \text{ with } n_{j} = \text{entier}[x 2^{-j} + \frac{1}{2}] \text{ and } x \in \mathbb{R},$$  \hspace{1cm} (3.A.4)

then

$$0 \leq x_{j} - x \leq \frac{1}{2} 2^{j}$$

and

$$\lim_{j \to -\infty} x_{j} = x.$$
Under the assumption that $\phi$ is left continuous this implies that
\[
\lim_{j \to -\infty} \phi\left(\frac{1}{2}x_j\right) = \phi\left(\frac{1}{2}x\right)
\]
and
\[
\lim_{j \to -\infty} \phi(x_j - k) = \phi(x - k).
\]
It remains to be shown that
\[
\lim_{j \to -\infty} \sum_{k=-\infty}^{\infty} h_j(k)\phi(x_j - k) - h(k)\phi(x - k) = 0. \tag{3.A.5}
\]
The limit and the sum in eq.(3.A.5) may be interchanged if
\[
S_N(j) = \sum_{k=-N}^{N} h_j(k)\phi(x_j - k) - h(k)\phi(x - k)
\]
converges uniformly. $S_N(j)$ converges uniformly if there exist $M_k$ and $K$ such that
\[
f_k(j) = |h_j(k)\phi(x_j - k) - h(k)\phi(x - k)| \leq M_k \quad \forall |k| \geq K, \forall j \leq 0,
\]
for which
\[
\sum_{|k| \geq K} M_k \text{ converges (Rudin[27])}.
\]
It is obvious that
\[
f_k(j) \leq |h_j(k)||\phi(x_j - k)| + |h(k)||\phi(x - k)|.
\]
Eq.(3.3.4) states that there exists a $K_1$ such that, for $|k| \geq K_1$,
\[
|\phi(x - k)| \leq \frac{M}{|k - x|^\alpha},
\]
\[
|\phi(x_j - k)| \leq \frac{M}{|k - x_j|^\alpha}.
\]
with $\alpha > \frac{2}{3}$.
Using eq.(3.3.4), eq.(3.6.2) and eq.(3.6.3), it can be shown that there exists a \( K_2 \) and \( M_1 \) such that, for \( |k| \geq K_2 \),

\[
\begin{align*}
|h_j(k)| &< \frac{M_2}{|k|^{\alpha}} , \\
|h(k)| &< \frac{M_2}{|k|^{\alpha}} ,
\end{align*}
\]

if \( \alpha > 1 \),

\[
\begin{align*}
|h_j(k)| &< \frac{M_2}{|k|^{\alpha-\epsilon}} , \\
|h(k)| &< \frac{M_2}{|k|^{\alpha-\epsilon}} ,
\end{align*}
\]

if \( \alpha = 1 \), for any \( \epsilon > 0 \),

and

\[
\begin{align*}
|h_j(k)| &< \frac{M_2}{|k|^{2\alpha - 1}} , \\
|h(k)| &< \frac{M_2}{|k|^{2\alpha - 1}} ,
\end{align*}
\]

if \( \frac{1}{2} < \alpha < 1 \).

Hence, we are able to find a \( M_k \) for which \( f_k(j) \leq M_k \) for \( |k| \) larger than a certain \( K \). It also follows that

\[
M_k = \begin{cases} 
O(|k|^{-2\alpha}) & \text{for } \alpha > 1 , \\
O(|k|^\epsilon^{-2\alpha}) & \text{for any } \epsilon > 0 , \text{for } \alpha = 1 , \\
O(|k|^{1-3\alpha}) & \text{for } \frac{1}{2} < \alpha < 1 .
\end{cases}
\]

This means that, for \( \alpha > \frac{2}{3} \),

\[
\sum_{|k| \geq K} M_k
\]

converges and, consequently, that \( S_N(j) \) converges uniformly. The limit and the sum in eq.(3.6.5) may therefore be interchanged, hence

\[
\lim_{j \to -\infty} \sum_{k=-\infty}^{\infty} h_j(k)\phi(x_j - k) - h(k)\phi(x - k) =
\]

\[
\sum_{k=-\infty}^{\infty} \lim_{j \to -\infty} h_j(k)\phi(x_j - k) - h(k)\phi(x - k) = 0
\]

This proves eq.(3.3.5). The proof for a right-continuous \( \phi \) is analogous but with \( n_j = \text{entier}[x2^{-j}] \) in eq.(3.6.4).
Eq. (3.3.6)

Fourier transformation of eq. (3.3.5) leads to eq. (3.3.11):

\[ \tilde{\phi}(2\omega) = H(\omega)\tilde{\phi}(\omega) . \]

This expression can be written as

\[ \tilde{\phi}(\omega) = \prod_{k=1}^{N} H(2^{-k}\omega)\tilde{\phi}(2^{-N}\omega) . \]

Because \( \tilde{\phi} \) is continuous in \( 0 \) and \( \tilde{\phi}(0) \neq 0 \) it follows that \( H(0) = 1 \) and that \( H \) is continuous in \( 0 \).

Eq. (3.3.7)

From eq. (3.3.5) it follows that

\[ \phi(n) = 2 \sum_{k=-\infty}^{\infty} h(k)\phi(2n - k) \]

or

\[ \phi(n) = 2 \sum_{k=-\infty}^{\infty} h(2n - k)\phi(k) . \]

When we define

\[ \Phi(\omega) = \sum_{n=-\infty}^{\infty} \phi(n) e^{-i\omega n} , \]

this expression can be written as

\[ \Phi(2\omega) = H(\omega)\Phi(\omega) + H(\omega + \pi)\Phi(\omega + \pi) . \]

The orthonormality of the \( \phi_k \)'s leads to

\[ \sum_{n=-\infty}^{\infty} \phi(n - k)\phi^*(n - m) = \begin{cases} 1 & m = k , \\ 0 & m \neq k , \end{cases} \]

which is equivalent with

\[ |\Phi(\omega)| = 1 \text{ or } \Phi(\omega) = e^{i\theta(\omega)} . \]
Consequently, we obtain eq. (3.3.7):

\[ H(\omega)e^{i\theta(\omega)} + H(\omega + \pi)e^{i\theta(\omega+\pi)} = e^{i\theta(2\omega)}, \]
where \( e^{i\theta(\omega)} = e^{i\theta(\omega+2\pi)} \).

Eq. (3.3.8)

Eq. (3.3.2) written in full, for \( j = -1 \), is

\[ 2^{-1} \sum_{n=-\infty}^{\infty} \phi(2^{-1}n - k)\phi^*(2^{-1}n - m) = \begin{cases} 1 & m = k, \\ 0 & m \neq k. \end{cases} \]

By eq. (3.3.5), we can write

\[ 2^{-1} \sum_{n=-\infty}^{\infty} \phi(2^{-1}n - k)\phi^*(2^{-1}n - m) = \]

\[ = 2 \sum_{p=-\infty}^{\infty} h(p - 2k) \sum_{l=-\infty}^{\infty} h^*(l - 2m) \sum_{n=-\infty}^{\infty} \phi(n - p)\phi^*(n - l). \]

The previous equation simplifies to

\[ 2^{-1} \sum_{n=-\infty}^{\infty} \phi(2^{-1}n - k)\phi^*(2^{-1}n - m) = 2 \sum_{p=-\infty}^{\infty} h(p - 2k)h^*(p - 2m). \]

This proves that

\[ 2 \sum_{p=-\infty}^{\infty} h(p - 2k)h^*(p - 2m) = \begin{cases} 1 & m = k, \\ 0 & m \neq k, \end{cases} \]

or

\[ 2 \sum_{p=-\infty}^{\infty} h(p - 2k)h^*(p) = \begin{cases} 1 & k = 0, \\ 0 & k \neq 0 \end{cases} \]

and this is equivalent with eq. (3.3.8). \( \square \)
3.A.2 Proof of Theorem 2

It will be proved that if φ satisfies eqs. (3.3.5) and (3.3.9), while H satisfies eq. (3.3.8), φ is a scaling function which means that eqs. (3.3.1) and (3.3.2) are satisfied.

Eq. (3.3.1) means that it should be possible to express $\phi_m^{j-1}$ as a linear combination of $\phi_k^j$'s.

$$\phi_m^{j-1}(n) = \sqrt{2^{j-1}} \phi(2^{j-1}n - m) = \sqrt{2} \sqrt{2^j} \sum_{k=-\infty}^{\infty} h(k) \phi(2^j n - 2m - k)$$

$$= \sqrt{2} \sqrt{2^j} \sum_{k=-\infty}^{\infty} h(k - 2m) \phi(2^j n - k)$$

Hence,

$$\phi_m^{j-1}(n) = \sqrt{2} \sum_{k=-\infty}^{\infty} h(k - 2m) \phi_k^j(n) .$$

Eq. (3.3.2) written in full is

$$2^j \sum_{n=-\infty}^{\infty} \phi(2^j n - k) \phi^*(2^j n - m) = \begin{cases} 1 & m = k , \\ 0 & m \neq k . \end{cases}$$

This expression will be denoted by $B_j$. It will be proved that $B_j \Rightarrow B_{j-1}$. Note that in this case $B_0$ holds by assumption.

By eq. (3.3.5), we can write

$$2^{j-1} \sum_{n=-\infty}^{\infty} \phi(2^{j-1} n - k) \phi^*(2^{j-1} n - m) =$$

$$2 \sum_{p=-\infty}^{\infty} h(p - 2k) \sum_{l=-\infty}^{\infty} h^*(l - 2m) 2^j \sum_{n=-\infty}^{\infty} \phi(2^j n - p) \phi^*(2^j n - l) .$$

If $B_j$ is true then the previous equations simplifies to

$$2^{j-1} \sum_{n=-\infty}^{\infty} \phi(2^{j-1} n - k) \phi^*(2^{j-1} n - m) = 2 \sum_{p=-\infty}^{\infty} h(p - 2k) h^*(p - 2m) .$$
From eq.(3.3.8) we can obtain that this expression becomes
\[ 2^{i-1} \sum_{n=-\infty}^{\infty} \phi(2^{i-1}n - k)\phi^*(2^{i-1}n - m) = \begin{cases} 1 & m = k, \\ 0 & m \neq k, \end{cases} \]
which is \( B_{j-1} \). □

### 3.A.3 Proof of Theorem 3

What has to be proved is the orthonormality of the \( \phi_k^0 \)'s. The remaining of the theorem follows from Theorem 2.

Eq.(3.3.2) with \( j = 0 \) is
\[ \sum_{n=-\infty}^{\infty} \phi(n - k)\phi^*(n - m) = \begin{cases} 1 & m = k, \\ 0 & m \neq k. \end{cases} \]

If \( \tilde{\phi} \in L^1(\mathbb{R}) \), which implies that \( \phi \) is continuous in every point \( x \in \mathbb{Z} \), we obtain, using the Poisson sum-formula, the following equation for \( \tilde{\phi} \):
\[ \sum_{k=-\infty}^{\infty} \tilde{\phi}(\omega + 2k\pi) = e^{i\theta(\omega)} \quad \forall \omega, \quad (3.3.6) \]
where clearly \( \theta \) should satisfy \( e^{i\theta(\omega)} = e^{i\theta(\omega+2\pi)}. \)

Next we define \( \tilde{\xi} \in L^1(\mathbb{R}) \) by
\[ \tilde{\xi}(\omega) = e^{-i\theta(\omega)}\tilde{\phi}(\omega). \]
Substitution of \( \tilde{\xi} \) in eq.(3.3.6) leads to
\[ \sum_{k=-\infty}^{\infty} \tilde{\xi}(\omega + 2k\pi) = 1 \quad \forall \omega. \quad (3.3.7) \]

Eq.(3.3.11) becomes
\[ \tilde{\xi}(2\omega) = e^{-i[\theta(2\omega)-\theta(\omega)]}H(\omega)\tilde{\xi}(\omega), \]
\[ = V(\omega)\tilde{\xi}(\omega) \]
and for \( V \) eqs.(3.3.7) and (3.3.8) become
\[ V(\omega) + V(\omega + \pi) = 1, \]
\[ |V(\omega)|^2 + |V(\omega + \pi)|^2 = 1. \]
Eq. (3.A.7) is equivalent with
\[
\xi(n) = \begin{cases} 
1 & n = 0, \\
0 & n \neq 0, \text{ with } n \text{ integer }.
\end{cases}
\] (3.A.8)

Eq. (3.A.8) is the property of \( \xi \) that will be proved.

Let \( P_k \) be defined by
\[
P_k(\omega) = \begin{cases} 
\prod_{p=1}^{k} V(2^{-p}\omega) \text{ for } |\omega| \leq 2^k\pi, \\
0 \text{ elsewhere.}
\end{cases}
\] (3.A.9)

Note that
\[
\lim_{k \to \infty} \prod_{p=1}^{k} V(2^{-p}\omega)
\]
exists because \(|V(\omega)| \leq 1\). Next we consider
\[
p_k(n) = \frac{1}{2\pi} \int_{-\infty}^{\infty} P_k(\omega)e^{i\omega n} \, d\omega
\]
\[
= \frac{1}{2\pi} \int_{-2^k\pi}^{2^k\pi} P_k(\omega)e^{i\omega n} \, d\omega
\]
\[
= \frac{1}{2\pi} \left[ \int_{-2^k\pi}^{0} P_k(\omega)e^{i\omega n} \, d\omega + \int_{0}^{2^k\pi} P_k(\omega)e^{i\omega n} \, d\omega \right].
\]
Changing the integration variable in the first integral and further simplification results in
\[
p_k(n) = \frac{1}{2\pi} 2^k\pi \int_{0}^{2^k\pi} (P_k(\omega) + P_k(\omega - 2^k\pi))e^{i\omega n} \, d\omega.
\]

By eq. (3.A.9) we have
\[
P_k(\omega - 2^k\pi) = \begin{cases} 
\prod_{p=1}^{k} V(2^{-p}(\omega - 2^k\pi - 2^{-p}\pi)) \text{ for } |\omega - 2^k\pi| \leq 2^k\pi, \\
0 \text{ elsewhere.}
\end{cases}
\]
or

\[ P_k(\omega - 2^k \pi) = \begin{cases} V(2^{-k}\omega - \pi) \prod_{p=1}^{k-1} V(2^{-p}\omega) & \text{for } |\omega - 2^k \pi| \leq 2^k \pi, \\ 0 & \text{elsewhere}. \end{cases} \]

Because \( V(\omega) + V(\omega + \pi) = 1 \) and \( V \) is \( 2\pi \)-periodic we find that

\[ P_k(\omega - 2^k \pi) = \begin{cases} (1 - V(2^{-k}\omega)) \prod_{p=1}^{k-1} V(2^{-p}\omega) & \text{for } |\omega - 2^k \pi| \leq 2^k \pi, \\ 0 & \text{elsewhere}. \end{cases} \]

Using this expression, \( p_k(n) \) becomes

\[ p_k(n) = \frac{1}{2\pi} \int_0^{2^k \pi} \prod_{p=1}^{k-1} V(2^{-p}\omega)e^{i\omega n} \, d\omega. \]

Using the periodicity of \( V \), this can be written as

\[ p_k(n) = \frac{1}{2\pi} \int_{-2^{k-1}\pi}^{2^{k-1}\pi} \prod_{p=1}^{k-1} V(2^{-p}\omega)e^{i\omega n} \, d\omega \]

\[ = \frac{1}{2\pi} \int_{-2^{k-1}\pi}^{2^{k-1}\pi} P_{k-1}(\omega)e^{i\omega n} \, d\omega \]

\[ = p_{k-1}(n). \]

Repetition of this procedure leads to

\[ p_k(n) = p_{k-1}(n) = \ldots = p_1(n). \]

\[ p_1(n) = \frac{1}{2\pi} \int_{-2\pi}^{2\pi} V(\frac{\omega}{2})e^{i\omega n} \, d\omega \]

\[ = \frac{1}{2\pi} \int_0^{2\pi} (V(\omega) + V(\omega + \pi)) e^{i\omega n} \, d\omega \]

\[ = \begin{cases} 1 & n = 0 \\ 0 & n \neq 0. \end{cases} \]
Eq.(3.3.10) ensures that

$$|V(2^{-k}\omega)| > 0 \ \forall \ \omega \in [-2^k\frac{\pi}{2}, 2^k\frac{\pi}{2}] .$$

Because $H$ is differentiable in 0 and $|H(0)| = 1$, we have

$$|V(\omega)| = |H(\omega)| = 1 + O(|\omega|)$$

in the neighborhood of 0. This implies that, for every $\omega \in [-\pi, \pi]$, there exist a $K$ such that

$$|\tilde{\xi}(\omega)| = \prod_{k=1}^{\infty} |V(2^{-k}\omega)|$$

$$= \exp\left(\sum_{k=1}^{\infty} \ln(|V(2^{-k}\omega)|)\right)$$

$$= \exp\left(\sum_{k=1}^{K-1} \ln(|V(2^{-k}\omega)|)\right) O(\exp(|\omega|2^{-(K-1)})) .$$

This means that

$$\exists C > 0 \text{ such that } |\tilde{\xi}(\omega)| \geq C \ \forall \ \omega \in [-\pi, \pi] .$$

Since $\tilde{\xi}(\omega) = P_k(\omega)\tilde{\xi}(2^{-k}\omega)$, it follows that

$$0 \leq |P_k(\omega)| \leq \frac{1}{C} |\tilde{\xi}(\omega)| \ \forall \ \omega .$$

The dominated convergence theorem (Weir[37]) now ensures that

$$\lim_{k \to \infty} \frac{1}{2\pi} \int_{-\infty}^{\infty} P_k(\omega)e^{i\omega n} \ d\omega = \frac{1}{2\pi} \int_{-\infty}^{\infty} \tilde{\xi}(\omega)e^{i\omega n} \ d\omega .$$

Thus,

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} \tilde{\xi}(\omega)e^{i\omega n} \ d\omega = \begin{cases} 1 & n = 0 , \\ 0 & n \neq 0 . \end{cases}$$

This proves that eq.(3.4.8), and therefore eq.(3.3.3), is satisfied. \Box
3.B Proof of Theorem 4

The necessity proof is analogous to the proof of Theorem 1 in Appendix 3.A.1. Derivation of the properties of $G$ is straightforward. Here, we will restrict ourselves to the sufficiency proof.

With eq.(3.4.7), calculations analogous to those in Appendix 3.A.2 show that 
\[ \{ \psi_k^j \}_{k=-\infty, \infty} \] 
is a set of orthonormal functions. Eq.(3.4.1) means that $\psi_m^{j-1}$ can be expressed as a linear combination of $\phi_k^j$'s. The proof is entirely analogous to that in Appendix 3.A.2 for $\phi_m^{j-1}$. Eq.(3.4.1) also means that 
\[ (\psi_k^j, \phi_m^j) = 0 \]
or
\[ 2^j \sum_{n=-\infty}^{\infty} \psi(2^j n - k)\phi^*(2^j n - m) = 0. \]

This expression will be denoted by $C_j$. For $B_j$ we take the same expression as in Appendix 3.A.2. It has already been proved that $B_j$ holds. Here, it will be shown that $B_j \Rightarrow C_{j-1}$. By eq.(3.3.5), we can write
\[ 2^{j-1} \sum_{n=-\infty}^{\infty} \psi(2^{j-1} n - k)\phi^*(2^{j-1} n - m) = \]
\[ 2 \sum_{p=-\infty}^{\infty} g(p - 2k) \sum_{l=-\infty}^{\infty} h^*(l - 2m) 2^j \sum_{n=-\infty}^{\infty} \phi(2^j n - p)\phi^*(2^j n - l). \]

Since $B_j$ is true, the previous equations simplifies to
\[ 2^{j-1} \sum_{n=-\infty}^{\infty} \psi(2^{j-1} n - k)\phi^*(2^{j-1} n - m) = 2 \sum_{p=-\infty}^{\infty} g(p - 2k)h^*(p - 2m). \]

By inverse Fourier transformation of eq.(3.4.4),
\[ G(\omega)H^*(\omega) + G(\omega + \pi)H^*(\omega + \pi) = 0, \]
we obtain that this expression can be written as
\[ 2^{j-1} \sum_{n=-\infty}^{\infty} \psi(2^{j-1} n - k)\phi^*(2^{j-1} n - m) = 0, \]
which is $C_{j-1}$.
The same eq.(3.4.1) also implies that $\phi^j_m$ can be expressed as a linear combination of $\phi^{j-1}_k$'s and $\psi^{j-1}_k$'s. Eqs.(3.3.5) and (3.4.2) lead to

$$
\phi(x - k) = 2 \sum_{n=-\infty}^{\infty} h(n - 2k)\phi(2x - n),
$$

and

$$
\psi(x - k) = 2 \sum_{n=-\infty}^{\infty} g(n - 2k)\psi(2x - n).
$$

From this we can obtain

$$
2 \sum_{n=-\infty}^{\infty} \phi(2x - n) \left[ \sum_{k=-\infty}^{\infty} h^*(-2k)h(n - 2k) + g^*(-2k)g(n - 2k) \right] =
\sum_{k=-\infty}^{\infty} h^*(-2k)\phi(x - k) + g^*(-2k)\psi(x - k)
$$

(3.B.1)

and

$$
2 \sum_{n=-\infty}^{\infty} \phi(2x - n) \left[ \sum_{k=-\infty}^{\infty} h^*(1-2k)h(n - 2k) + g^*(1-2k)g(n - 2k) \right] =
\sum_{k=-\infty}^{\infty} h^*(1-2k)\phi(x - k) + g^*(1-2k)\psi(x - k).
$$

(3.B.2)

The discrete-time Fourier transform of the term between brackets in eq.(3.B.1) is

$$
\frac{1}{2} \left[ H(\omega)[H^*(\omega) + H^*(\omega + \pi)] + G(\omega)[G^*(\omega) + G^*(\omega + \pi)] \right].
$$

The discrete-time Fourier transform of the term between brackets in eq.(3.B.2) is

$$
\frac{1}{2} \left[ H(\omega)[H^*(\omega) - H^*(\omega + \pi)] + G(\omega)[G^*(\omega) - G^*(\omega + \pi)] \right].
$$

From eqs.(3.3.8), (3.4.3) and (3.4.4), we can derive that both terms equal $\frac{1}{2}$. This can be shown as follows. By eq.(3.4.3),
\[ \frac{1}{2} [H(\omega)[H^*(\omega) \pm H^*(\omega + \pi)] + G(\omega)[G^*(\omega) \pm G^*(\omega + \pi)]] = \]
\[ \frac{1}{2} \pm \frac{1}{2} [H(\omega)H^*(\omega + \pi) + G(\omega)G^*(\omega + \pi)] . \]

After multiplication by \( G(\omega + \pi) \) the term between brackets becomes
\[ H(\omega)G(\omega + \pi)H^*(\omega + \pi) + G(\omega)|G(\omega + \pi)|^2 . \]

With eq.(3.4.4) we arrive at
\[ -G(\omega)|H(\omega)|^2 + G(\omega)|G(\omega + \pi)|^2 \]
and since \(|H(\omega)| = |G(\omega + \pi)|\) this expression equals 0. Consequently,
\[ \frac{1}{2} [H(\omega)[H^*(\omega) \pm H^*(\omega + \pi)] + G(\omega)[G^*(\omega) \pm G^*(\omega + \pi)]] = \frac{1}{2} \]
and therefore
\[ \sum_{k=-\infty}^{\infty} h^*(-2k)h(n-2k) + g^*(-2k)g(n-2k) = \begin{cases} 1 \ n = 0 , \\ 0 \ n \neq 0 \end{cases} \]
and
\[ \sum_{k=-\infty}^{\infty} h^*(1-2k)h(n-2k) + g^*(1-2k)g(n-2k) = \begin{cases} 1 \ n = 0 , \\ 0 \ n \neq 0 \end{cases} . \]

Hence, eq.(3.B.1) simplifies to
\[ \phi(2x) = \sum_{k=-\infty}^{\infty} h^*(-2k)\phi(x - k) + g^*(-2k)\psi(x - k) \]
and eq.(3.B.2) simplifies to
\[ \phi(2x) = \sum_{k=-\infty}^{\infty} h^*(1-2k)\phi(x - k) + g^*(1-2k)\psi(x - k) . \]

Substitution of \( x = 2^{j-1}n - \frac{1}{2}m \) in the former two equations results in
\[ \phi(2^j n - m) = \]
\[ \sum_{k=-\infty}^{\infty} h^*(-2k)\phi(2^{j-1}n - \frac{m}{2} - k) + g^*(-2k)\psi(2^{j-1}n - \frac{m}{2} - k) \quad (3.B.3) \]
and
\[
\phi(2^i n - m) = \sum_{k=-\infty}^{\infty} h^*(1 - 2k)\phi(2^{i-1} - \frac{m}{2} - k) + g^*(1 - 2k)\psi(2^{i-1} - \frac{m}{2} - k). \tag{3.B.4}
\]

Changing the summation variable in eq.(3.B.3) in \( k_1 = k + \frac{m}{2} \) for \( m \) even and changing the summation variable in eq.(3.B.4) in \( k_1 = k + \frac{m}{2} + \frac{1}{2} \) for \( m \) odd, leads to a linear relation between \( \phi_j^i \), the \( \phi_k^{j-1} \)'s and the \( \psi_k^{j-1} \)'s that is given by
\[
\phi_j^i = \sqrt{2} \sum_{k=-\infty}^{\infty} h^*(m - 2k) \phi_k^{j-1} + g^*(m - 2k) \psi_k^{j-1}.
\]

\[\square\]

\section*{3.C Miscellaneous}

\textbf{Decay rate of} \( h \)

Suppose that \( H \in C^q \).
\[
h(n) = \frac{1}{2\pi} \int_{-\pi}^{\pi} H(\omega)e^{i\omega n} \, d\omega.
\]

Partial integration results in
\[
h(n) = -\frac{1}{2\pi in} \int_{0}^{2\pi} \frac{dH(\omega)}{d\omega} e^{i\omega n} \, d\omega,
\]
\[
= \frac{1}{2\pi \left( \frac{i}{n} \right)^{q+1}} \int_{0}^{2\pi} \frac{d^{q+1} H}{d\omega^{q+1}}(\omega)e^{i\omega n} \, d\omega, \ n \neq 0.
\]

\( \frac{d^{q+1} H}{d\omega^{q+1}}(\omega) \) may have some jump-discontinuities but still
\[
|\int_{0}^{2\pi} \frac{d^{q+1} H}{d\omega^{q+1}}(\omega)e^{i\omega n} \, d\omega| = o(1) \text{ as } |n| \to \infty.
\]

This shows that
\[
|h(n)| = o\left( \frac{1}{|n|^{q+1}} \right) \text{ as } |n| \to \infty. \quad \square
\]
Taylor expansion of $H$

If $H \in C^q$ and if

$$\frac{d^n H}{d\omega^n}(\pi) = 0 \text{ for } 0 \leq n \leq q - 1$$

then Taylor expansion around $\omega = \pi$ yields

$$H(\omega) = \frac{d^q H}{d\omega^q}(\xi(\omega)) \frac{(\omega - \pi)^q}{q!}.$$

with $\xi(\omega)$ in between $\pi$ and $\omega$. This allows us to write

$$H(\omega) = (\cos(\frac{\omega}{2}))^q T(\omega),$$

with

$$|T(\omega)| < \infty.$$  

Note that in the case we have

$$H(\omega) + H(\omega + \pi) = 1;$$

we do obtain $\frac{d^q H}{d\omega^q}(\pi) = 0$ but this does not render continuity of $\frac{d^{q+1} H}{d\omega^{q+1}}$. 
Chapter 4

Multiscale segmentation of well logs

4.1 Introduction

Well logs often clearly exhibit some global character with a superposition of local features. This observation calls for processing methods that can reveal this behavior to improve analysis.

Segmentation of well logs is often performed by detecting edges and interpreting them as boundaries between the different geological units. Edge detection can be done by smoothing and then differentiating the log. Where the derivative of the smoothed log exceeds a certain threshold, an edge is detected. In this way, however, the characteristic multiscale behavior of the log is not revealed.

In this chapter a multiscale segmentation of well logs is introduced, that does reveal this characteristic behavior. The segmentation begins with a multiscale transform of the well log. This transform results in a separation of local, high-frequent, and global, low-frequent, behavior. The segmentation is based on the behavior of edges at various scales. The next step is to combine results obtained at different scales. Processing can start at a coarse scale, where noise has been reduced considerably. The results obtained there can be used to descend to finer scales.

Theory on multiscale edge detection is provided in the first section. In the second section some possible ways of segmentation, with examples, are considered.

---

4 A part of this chapter has appeared as a report of the Faculty of Technical Mathematics and Informatics and as a paper in Mathematical Geology (Vermeer and Alkemade[31],[32]).
4.2 Multiscale edge detection

4.2.1 General

Edge detection is of great importance in signal analysis. Edges are associated with extrema of the first derivative and consequently with zero crossings of the second derivative (see fig.4.2.1).

Usually a signal is smoothed before differentiation in order to increase the signal-to-noise ratio. However, smoothing results in a loss of localization of the edge positions because of interference between edges with distances apart of the order of the width of the smoothing filter (Canny[6]; Wilson and Spann[38]).

Multiscale edge detection aims to restore the localization by gradually descending from coarse scales to fine scales.

4.2.2 Scale space

4.2.2.1 Definition and example

The behavior of the zero crossings of the second derivative across different scales was first investigated by Witkin[39]. We determine the second derivative of the multiscale transform (eq.(2.2.2)) of a function \( f \) with respect to the variable \( x \),

\[
\frac{\partial^2}{\partial x^2} M(f)(x, \sigma) = \frac{1}{\sigma^3} \int_{-\infty}^{\infty} f(t) g^{(2)}(\frac{x-t}{\sigma}) \, dt \tag{4.2.1}
\]

It is assumed that the derivatives of \( g \) that are used in the sequel are all such that their convolution with \( f \) renders a finite result (see eq.(2.1.2)).

Figure 4.2.1: Example of an edge with its first and second derivative
Definition
The scale space of a signal $f$ with respect to the second derivative is defined as
\[ I(f) = \{(x, \sigma) \mid \frac{\partial^2}{\partial x^2} M(f)(x, \sigma) = 0, \frac{\partial^3}{\partial x^3} M(f)(x, \sigma) \neq 0, \sigma > 0\} \] (4.2.2)
(see Witkin[39]; Babaud et al.[2]; Yuille and Poggio[41] and Koenderink[19]).

In eq.(4.2.2) inflection points of $\frac{\partial^2}{\partial x^2} M(f)(x, \sigma)$ are excluded from the scale space, since they do not correspond to edges.

We now formulate one additional condition on the filter $g$ in the multiscale transform. The zero crossings $x(\sigma)$ of an ideal step edge,
\[ u(x) = \begin{cases} 1 & x \geq 0, \\ 0 & x < 0, \end{cases} \]
are solutions of
\[ \frac{\partial^2}{\partial x^2} M(u)(x, \sigma) = \frac{1}{\sigma^2} g^{(1)}\left(\frac{x}{\sigma}\right) = 0, \quad \frac{1}{\sigma^3} g^{(2)}\left(\frac{x}{\sigma}\right) \neq 0. \]
In this case it is reasonable to require that only one zero crossing exists at each scale and that it is located in the origin. This requirement is fulfilled if $g$ has only one extremum, which is located in the origin.

Remark 1
Zero crossings of other derivatives of the multiscale transform, may also be examined. This allows to detect features other than edges in a signal, such as peaks and corners.

Remark 2
The second derivative of $g$ is an analyzing wavelet for the wavelet transform. Except for a factor $\sigma^2 \sqrt{\sigma}$ the transform in eq.(4.2.1) is therefore a wavelet transform of $f$. Determination of the scale space of $f$ is equivalent to determination of the zero crossings of its wavelet transform (Vermeer and Alkemade[33]).
In fig.4.3.1 a gamma-ray log and its corresponding scale space are given as an example. The applied filter is the Gaussian,

$$g(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}$$

The scale space consists of curves, beginning at the finest scale (\(\sigma = 0\)). The curves may vanish at a certain maximum scale. In this picture curves vanish only in pairs. Such a pair of curves will be called a zero-crossing contour. It is obvious that the global edges can be determined at a coarse scale and that their exact locations can be found by coarse-to-fine tracking.

4.2.2.2 The non-creation property

It would be desirable if no zero crossings were created for increasing scale \(\sigma\). The first reason is that we do not want the amount of data to increase as scale becomes coarser. The second reason is that we want to prevent that segments in the signal split up as scale becomes coarser, because that would violate a sense of causality.

Here only stable zero crossings will be considered (Yuille and Poggio[41]). These can only disappear in pairs in the \((x, \sigma)\)-plane.

**Theorem**  The only filter \(g\) in the multiscale transform that does not allow creation of zero crossings for increasing scale \(\sigma\) is the Gaussian.

(Babaud et al.[2] and Yuille and Poggio[41])

We note that this property is valid for the scale space with respect to any derivative.

**Proof**

For convenience \(\frac{\partial^2}{\partial x^2} M(f)(x, \sigma)\) will be denoted by \(E(x, \sigma)\).

The zero crossings are solutions of

$$E(x, \sigma) = 0 \quad , \quad \frac{\partial}{\partial x} E(x, \sigma) \neq 0 .$$

No zero crossings are created for increasing \(\sigma\) if a zero-crossing contour can be closed above but never below. Hence, if at points \((x, \sigma)\) where
\frac{d\sigma}{dx} = 0 ,

it can be guaranteed that
\frac{d^2\sigma}{dx^2} < 0 . \tag{4.2.3}

Along a zero-crossing curve
\[ E = 0 \text{ and consequently } \frac{d^m E}{dx^m} = 0 \text{ for } m \geq 1 . \]

This leads to
\frac{dE}{dx} = \frac{\partial E}{\partial x} + \frac{\partial E}{\partial \sigma} \frac{d\sigma}{dx} = 0

and therefrom to
\frac{d\sigma}{dx} = - \frac{\partial E}{\partial x} \bigg/ \frac{\partial E}{\partial \sigma} . \tag{4.2.4}

Calculation of the second total derivative with respect to \( x \) gives
\[ \frac{\partial^2 E}{\partial x^2} + \frac{\partial E}{\partial \sigma} \frac{d^2\sigma}{dx^2} = 0 \text{ at points where } \frac{d\sigma}{dx} = 0 . \]

This shows that
\frac{d^2\sigma}{dx^2} = - \frac{\partial^2 E}{\partial x^2} \bigg/ \frac{\partial E}{\partial \sigma} .

Thus, eq.(4.2.3) holds if
\[ \frac{\partial^2 E}{\partial x^2} \bigg/ \frac{\partial E}{\partial \sigma} > 0 . \tag{4.2.5} \]

**Sufficiency**

The heat equation can be written as
\[ \frac{\partial^2 E}{\partial x^2} = \frac{1}{\sigma} \frac{\partial E}{\partial \sigma} . \]

Note that by substitution of \( t = \frac{\sigma^2}{2} \), we obtain the standard heat equation.

The Gaussian (eq.(2.2.3)) is the Green's function of the heat equation. If the Gaussian is used as filter, \( M(f) \) is a superposition of Gaussians and therefore
satisfies the heat equation, as do all its derivatives. Hence, the inequality in eq.(4.2.5) will hold.

Necessity
In Babaud et al.[2] and Yuille and Poggio[41] the necessity proof is based on the construction of an example that violates eq.(4.2.5) unless the applied filter is the Gaussian. Here, we will use a different approach.

Let us equate the quotient in eq.(4.2.5) to a function \( h = h(x, \sigma) \), or

\[
\frac{\partial^2 E}{\partial x^2}(x, \sigma) = h(x, \sigma) \frac{\partial E}{\partial \sigma}(x, \sigma) .
\]

(4.2.6)

The function \( h \) should be positive at points of inflection of the zero-crossing contours. When we take \( f \) to be a periodic function,

\[
f(x) = f(x + nT) \quad \forall \ n \in \mathbb{Z} ,
\]

it follows that

\[
E(x, \sigma) = E(x + nT, \sigma) \quad \forall \ n \in \mathbb{Z} ,
\]

which is also true for all partial derivatives of \( E \) and by eq.(4.2.6) also for \( h \). Hence,

\[
h(x, \sigma) = h(x + nT, \sigma) \quad \forall \ n \in \mathbb{Z} .
\]

However, since the period \( T \) can be chosen arbitrarily, \( h \) has to be independent of \( x \), leaving for eq.(4.2.6),

\[
\frac{\partial^2 E}{\partial x^2}(x, \sigma) = h(\sigma) \frac{\partial E}{\partial \sigma}(x, \sigma) .
\]

(4.2.7)

We know that \( h \) should be positive at points of inflection of the zero-crossing contours. Since it has be shown that \( h \) is independent of \( f \) and \( x \), this implies that

\[
h(\sigma) > 0 \quad \forall \ \sigma > 0 .
\]

By substitution of

\[
t = \int_0^\sigma \frac{1}{h(s)} \, ds
\]
in eq.(4.2.7) we obtain the standard heat equation of which the Green’s function is
\[ g_{\text{heat}}(x,t) = \frac{1}{2\sqrt{t\pi}} e^{-\frac{x^2}{4t}}. \]

By inspection of the definition of the multiscale transform in eq.(2.2.2), it becomes clear that
\[ t = \frac{a}{2} \sigma^2 \quad \text{for some } a > 0 \]

and consequently
\[ h(\sigma) = \frac{1}{a\sigma} \quad \text{for some } a > 0. \]

This proves that just by inspection of the properties of the quotient in eq.(4.2.5), the Gaussian emerges as the only possible filter. The factor \( a \) is only a scaling factor and can be set at the value 1 without any loss of generality. \( \square \)

### 4.2.3 Numerical aspects

#### 4.2.3.1 Calculation of the multiscale transform

In practice a function \( f \) will only be given in a number of data points
\[ f_n = f(n) , \quad n = 0, \ldots, N - 1. \]

We assume that \( f \) can be approximated by a sum of Gaussians,
\[ \hat{f}(x) = \sum_{n=0}^{N-1} \alpha_n \ g_{\sigma_0}(x - n) , \quad \sigma_0 > 0 , \quad (4.2.8) \]

in which \( \{ \alpha_n \}_{n=0,\ldots,N-1} \) and \( \sigma_0 \) are yet to be determined. For convenience, we have adopted the notation
\[ g_{\sigma}(x) = e^{-\frac{1}{2}(\frac{x}{\sigma})^2}. \]

The advantage of eq.(4.2.8) becomes clear in the next theorem.
Theorem  The multiscale transform of $f$ at scale $\sigma$ is given by
\[
M(\hat{f})(x, \sigma) = \frac{\sigma_0}{(\sigma_0^2 + \sigma^2)^{1/2}} \sum_{n=0}^{N-1} \alpha_n g_{(\sigma_0^2 + \sigma^2)^{1/2}}(x - n). \tag{4.2.9}
\]

With eq.(4.2.9) we have a formula for $M(\hat{f})$ that has the same form for all scales $\sigma$. Once the coefficients $\alpha_n$ are known, $M(\hat{f})$ and all its derivatives can be calculated.

Proof
\[
M(\hat{f})(x, \sigma) = \frac{1}{\sigma \sqrt{2\pi}} \int_{-\infty}^{\infty} \hat{f}(t) g_\sigma(x - t) \, dt \\
= \sum_{n=0}^{N-1} \alpha_n \frac{1}{\sigma \sqrt{2\pi}} \int_{-\infty}^{\infty} g_{\sigma_0}(t - n) g_\sigma(x - t) \, dt. \tag{4.2.10}
\]

With Parseval's theorem,
\[
\frac{1}{\sigma \sqrt{2\pi}} \int_{-\infty}^{\infty} g_{\sigma_0}(t - n) g_\sigma(x - t) \, dt = \frac{1}{\sigma (2\pi)^{3/2}} \int_{-\infty}^{\infty} \hat{g}_{\sigma_0}(\omega) \hat{g}_\sigma(\omega) \, e^{i\omega(x-n)} \, d\omega.
\]

Because
\[
\hat{g}_\sigma(\omega) = \sigma \sqrt{2\pi} \, e^{-\frac{1}{2} \sigma^2 \omega^2},
\]
it follows that
\[
\frac{1}{\sigma \sqrt{2\pi}} \int_{-\infty}^{\infty} g_{\sigma_0}(t - n) g_\sigma(x - t) \, dt = \frac{\sigma_0}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{2} (\sigma_0^2 + \sigma^2) \omega^2} \, e^{i\omega(x-n)} \, d\omega.
\]
\[
= \frac{\sigma_0}{(\sigma_0^2 + \sigma^2)^{1/2}} g_{(\sigma_0^2 + \sigma^2)^{1/2}}(x - n).
\]

Substitution of this expression in eq.(4.2.10) yields eq.(4.2.9). $\square$
Theorem. For given \( \sigma_0 \), the coefficients \( \alpha_n \), for which eq. (4.2.8) is an interpolation formula, are given by

\[
\alpha_n = \frac{1}{2\pi} \int_0^{2\pi} \frac{F(\omega)}{\theta_{\sigma_0}(\omega)} e^{i\omega n} \, d\omega , \quad n = 0, \ldots, N - 1 .
\] (4.2.11)

where \( F \) is the discrete-time Fourier transform of \( f(n) \) and where

\[
\theta_{\sigma}(\omega) = \sigma \sqrt{2\pi} \sum_{k=-\infty}^{\infty} e^{-\frac{1}{2}\sigma^2(\omega + 2k\pi)^2} .
\]

Proof.
From eq. (4.2.8) we obtain the interpolation condition,

\[
f(n) = \sum_{k=0}^{N-1} \alpha_k g_{\sigma_0}(n-k) .
\]

Discrete-time Fourier transformation of both sides of this expression results in

\[
F(\omega) = A(\omega) \sum_{n=-\infty}^{\infty} g_{\sigma_0}(n) e^{-i\omega n} .
\]

We introduce

\[
\theta_{\sigma}(\omega) = \sum_{n=-\infty}^{\infty} g_{\sigma}(n) e^{-i\omega n} = \sum_{k=-\infty}^{\infty} \tilde{g}_{\sigma}(\omega + 2k\pi) = \sigma \sqrt{2\pi} \sum_{k=-\infty}^{\infty} e^{-\frac{1}{2}\sigma^2(\omega + 2k\pi)^2} \] (4.2.12)

The previous steps can be made by applying the Poisson summation formula. Hence,

\[
F(\omega) = A(\omega) \theta_{\sigma_0}(\omega) .
\]

In eq. (4.2.12) it can be seen that

\[
\theta_{\sigma}(\omega) > 0 \quad \forall \omega .
\]

Moreover, it can be shown that
\[ 0 < \sigma \sqrt{2\pi} \ m_\sigma \leq \theta_\sigma(\omega) \leq \sigma \sqrt{2\pi} \ M_\sigma , \]

in which

\[ M_\sigma = 1 + \sum_{k=1}^{\infty} \left( e^{-\frac{1}{2}\sigma^2(2k\pi)^2} + e^{-\frac{1}{2}\sigma^2((2k-1)\pi)^2} \right) \]

and

\[ m_\sigma = e^{-\frac{1}{2}\sigma^2\pi^2} + \sum_{k=1}^{\infty} \left( e^{-\frac{1}{2}\sigma^2(2k\pi)^2} + e^{-\frac{1}{2}\sigma^2((2k+1)\pi)^2} \right). \quad (4.2.13) \]

Now that we have established that \( \theta_\sigma \) is essentially positive we may write

\[ A(\omega) = \frac{F(\omega)}{\theta_{\sigma_0}(\omega)}. \quad (4.2.14) \]

Eq. (4.2.11) is the inverse Fourier transform of this expression. \( \square \)

If eq. (4.2.8) is an interpolation formula we can write

\[ \hat{f}(x) = \sum_{n=0}^{N-1} f_n \ \phi_{\sigma_0}(x - n) , \text{ for any } \sigma_0 > 0, \quad (4.2.15) \]

where

\[ \phi_{\sigma_0}(x) = \begin{cases} 1 & \text{if } x = 0, \\ 0 & \text{if } x \in \mathbb{Z} \setminus \{0\}. \end{cases} \]

**Theorem** The interpolation function \( \phi_{\sigma_0} \) can be defined through its Fourier transform,

\[ \tilde{\phi}_{\sigma_0}(\omega) = \frac{\tilde{g}_{\sigma_0}(\omega)}{\theta_{\sigma_0}(\omega)} = \frac{e^{-\frac{1}{2}\sigma_0^2\omega^2}}{\sum_{k=-\infty}^{\infty} e^{-\frac{1}{2}\sigma_0^2(\omega+2k\pi)^2}}. \quad (4.2.16) \]

**Proof**
The Fourier transform of \( \phi_{\sigma_0} \) follows by Fourier transformation of eqs. (4.2.8), (4.2.15) and substitution of eq. (4.2.14). Because \( \tilde{\phi}_{\sigma_0} \) has exponential decay it uniquely determines \( \phi_{\sigma_0} \). \( \square \)
Since eq.(4.2.15) is valid for any positive value of $\sigma_0$, eq.(4.2.16) defines a class of interpolation functions. This is a degree of freedom that requires some attention. If we take $\sigma_0$ to be very small $\tilde{f}$ becomes a sum of very narrow Gaussians. This implies that, although $\tilde{f}$ attains the sample values in the sample points, it will be almost zero in between the sample points. Obviously, this is undesirable. Another interesting effect occurs if we allow $\sigma_0$ to tend to infinity.

**Theorem** The limit of $\tilde{\phi}_{\sigma_0}$ as $\sigma_0$ tends to infinity is given by

$$
\lim_{\sigma_0 \to \infty} \tilde{\phi}_{\sigma_0}(\omega) = \begin{cases} 
1 & -\pi < \omega < \pi , \\
\frac{1}{2} & \omega = \pm \pi , \\
0 & \text{elsewhere}.
\end{cases}
$$

(4.2.17)

This limit holds pointwise as well as in any $L^p$ norm.

**Proof**

By eq.(4.2.16),

$$
\tilde{\phi}_{\sigma_0}(\omega) = \frac{e^{-\frac{1}{2}\sigma_0^2 \omega^2}}{\sum_{k=-\infty}^{\infty} e^{-\frac{1}{2}\sigma_0^2 (w+2k\pi)^2}} = \frac{1}{1 + \sum_{k \neq 0} e^{-2\sigma_0^2 k \pi (\omega + k \pi)}}.
$$

(4.2.18)

Consider the denominator of this expression. The sign behavior of the exponent of $e$ in eq.(4.2.18) as a function of $k$ is given by

- For $\omega < 0$,

  $$
  -2\sigma_0^2 k \pi (\omega + k \pi) \begin{cases} 
  > 0 & \text{for } 0 < k < -\frac{\omega}{\pi} , \\
  = 0 & \text{if } k = 0 \text{ or } k = -\frac{\omega}{\pi} , \\
  < 0 & \text{elsewhere}.
  \end{cases}
  $$

- For $\omega > 0$,

  $$
  -2\sigma_0^2 k \pi (\omega + k \pi) \begin{cases} 
  > 0 & \text{for } -\frac{\omega}{\pi} < k < 0 , \\
  = 0 & \text{if } k = 0 \text{ or } k = -\frac{\omega}{\pi} , \\
  < 0 & \text{elsewhere}.
  \end{cases}
  $$
- For $\omega = 0$,

\[-2\sigma_0^2 k\pi (\omega + k\pi) < 0.\]

This results in the following conclusions for the behavior of the denominator in eq.(4.2.18).

- If $|\frac{\omega}{\pi}| > 1$ then the denominator can be written as

\[1 + e^{-2\sigma_0^2 \pi (\omega + \pi)} + e^{2\sigma_0^2 \pi (\omega - \pi)} + \sum_{k \neq -1,0,1} e^{-2\sigma_0^2 k\pi (\omega + k\pi)}\]

This expression tends to infinity as $\sigma_0$ tends to infinity.

- If $|\frac{\omega}{\pi}| < 1$ then the denominator

\[1 + \sum_{k \neq 0} e^{-2\sigma_0^2 k\pi (\omega + k\pi)} \quad (4.2.19)\]

tends to 1 as $\sigma_0$ tends to infinity because the exponent of the exponential function is negative for all $k \neq 0$.

- If $\omega = \pm \pi$ then the denominator becomes

\[1 + \sum_{k \neq 0} e^{-2\sigma_0^2 k\pi^2 (k \pm 1)} \quad (4.2.20)\]

This expression tends to 2 as $\sigma_0$ tends to infinity.

The pointwise convergence is hereby established. It remains to be shown that

\[\lim_{\sigma_0 \to \infty} \left( \int_{-\infty}^{\infty} |\tilde{\phi}_\infty(\omega) - \tilde{\phi}_{\sigma_0}(\omega)|^p \, d\omega \right)^{\frac{1}{p}} = 0 \quad (4.2.21)\]

By using eqs.(4.2.13), (4.2.18), (4.2.19) and (4.2.20) it follows that

\[\tilde{\phi}_{\sigma_0}(\omega) \leq \begin{cases} 1 & |\omega| \leq \pi, \\ e^{-\frac{1}{2} \sigma_0^2 (\omega^2 - \pi^2)} & |\omega| > \pi. \end{cases} \quad (4.2.22)\]
Next, we write

$$|\tilde{\phi}_\infty(\omega) - \tilde{\phi}_{\sigma_0}(\omega)|^p \leq |\tilde{\phi}_\infty(\omega) + \tilde{\phi}_{\sigma_0}(\omega)|^p$$

For $\eta$ arbitrarily positive and by eq.(4.2.22), a function $\mu \in L^1(\mathbb{R})$ can be found such that

$$|\tilde{\phi}_\infty(\omega) + \tilde{\phi}_{\sigma_0}(\omega)|^p \leq \mu(\omega) \quad \forall \sigma_0 > \eta .$$

By the dominated convergence theorem (Weir[37]) the limit and integral in eq.(4.2.21) may now be interchanged and convergence in $L^p$ sense has been proved. □

The inverse Fourier transform of $\tilde{\phi}_\infty$ is the sinc-function,

$$\text{sinc}(x) = \frac{\sin(\pi x)}{\pi x} .$$

Interpolation with the sinc-function has disadvantages as well because it is known that if a sampled ideal edge is interpolated using sinc-functions the resulting function exhibits Gibbs-alike phenomena near the location of the edge.

For the interpolation we will use a value of $\sigma_0$ for which the interpolating function $\phi_{\sigma_0}$ is as smooth as possible in some sense. The smoothness of the interpolating function will be defined by the effective width of its Fourier transform. The effective width of a function has been defined in the definitions. We seek to find the value of $\sigma_0$ for which $\Delta \omega$ is minimal. For $\tilde{\phi}_\infty$ the effective width can be calculated straightforwardly using eq.(4.2.17),

$$\Delta \omega = \frac{\pi}{\sqrt{3}} \approx 1.814 .$$

For other $\sigma_0$ the effective width has been determined numerically. The results are depicted in fig.4.2.2. It can be observed that $\Delta \omega$ has a minimum and numerically we have found that this minimum is attained for

$$\sigma_0 \approx 0.613 \text{ with } \Delta \omega \approx 1.690 .$$

The corresponding $\phi_{0.613}$ is depicted in fig.4.2.3. The same $\phi_{0.613}$ is depicted in fig.4.2.4 together with $\phi_{0.4}$ and $\phi_{0.8}$. 
Figure 4.2.2: The effective width ($\Delta \omega$) of $\tilde{f}_{\sigma_0}$ as a function of $\sigma_0$.

Figure 4.2.3: The interpolation function $\phi_{\sigma_0}$ with $\sigma_0 = 0.613$. 
4.2.3.2 Tracking of zero crossings

In this section a qualitative description of a method for the tracking of zero crossings in the scale space (eq.(4.2.2)) is given.

The problem of tracking the zero crossings is a consequence of the discretization in $\sigma$ and amounts up to the connection of each zero crossing at scale $\sigma$ to a zero crossing at a smaller scale or conversely to a zero crossing at a larger scale, provided the latter is possible.

This connection can be made as follows. Suppose that $x(\sigma)$ is a zero crossing at scale $\sigma$. We suggest to calculate the corresponding $x(\sigma \pm \Delta \sigma)$ using an iterative scheme for the determination of zeros which requires only one start value. In formula, $x(\sigma \pm \Delta \sigma)$ is determined by a scheme of the form

$$x_{k+1} = S(x_k, M(f), \frac{\partial}{\partial x} M(f), ..) \ , \text{with} \ x_0 = x(\sigma) ,$$

for example a Newton-Raphson scheme

$$x_{k+1} = x_k + \frac{\partial^2}{\partial x^2} M(f)(x_k, \sigma \pm \Delta \sigma) \left/ \frac{\partial^3}{\partial x^3} M(f)(x_k, \sigma \pm \Delta \sigma) \right. .$$

Note that the results of the previous section allow us to calculate the multiscale transform and its derivatives in any point in the $(x, \sigma)$-plane. This means that
the evaluation of a function $S$ in the iterative scheme is possible. The step in scale of $\Delta \sigma$ must be small enough to ensure convergence from $x(\sigma)$ to the correct $x(\sigma \pm \Delta \sigma)$. Precaution has to be taken to deal with the situation of an inflection point, i.e. the top of a zero crossing contour.

In the preceding we have assumed that there actually exists a zero crossing at scale $\sigma \pm \Delta \sigma$ that corresponds to the known zero crossing at scale $\sigma$. However, when stepping in the direction of increasing scale this need not be the case, since we know that zero crossings may vanish for increasing scale. The iterative scheme will probably still come up with a solution. It is therefore necessary to compare the solutions that have been found to the initial solution at a smaller scale and more important to solutions obtained for other zero crossings.

A top-down approach induces another problem. If every zero crossing at scale $\sigma$ has been connected to a zero crossing at scale $\sigma - \Delta \sigma$, there can be new zero crossings that remain to be found. This implies that the entire signal at that particular scale has to be scanned in search for additional zero crossings, which will be a cumbersome task. If however one is only interested in tracking the zero crossings that extend at least to a certain scale $\sigma_{\text{min}}$ the amount of work can be reduced.

It is for this reason that in general the top-down approach is preferable.
4.3 Multiscale segmentation

4.3.1 General

In this section the ideas and theory of multiscale signal processing are applied to the segmentation of well logs. The need for multiscale signal processing in well log analysis has already been expressed in Collins and Doveton[7]. Well logs exhibit characteristic behavior on multiple scales. Therefore, multiscale edge detection may produce a segmentation that fits the general nature of the log.

4.3.2 Modified scale space

A common segmentation of a well log \( f \) produces

\[
\text{\( k \) edges : } x_1, \ldots, x_k \\
\text{and} \\
\text{\( (k - 1) \) segments : } I_j = [x_j, x_{j+1}) , \quad j = 1, \ldots, k - 1.
\]

The scale space yields

\[
\text{\( k \) edges at scale } \sigma : \quad x^{(\sigma)}_1, \ldots, x^{(\sigma)}_k, \quad \forall \sigma \geq 0 \\
\text{and} \\
\text{\( (k - 1) \) segments : } I^{(\sigma)}_j = [x^{(\sigma)}_j, x^{(\sigma)}_{j+1}) , \quad j = 1, \ldots, k - 1 \quad \forall \sigma \geq 0
\]

Note that \( k^{(\sigma_1)} \leq k^{(\sigma_2)} \) if \( \sigma_1 \geq \sigma_2 \), in view of the conditions we have imposed on the filter \( g \). The \( x^{(\sigma)}_j \) only represent approximations of the correct edge locations, which is a result of the smoothing. The localization can be restored by coarse-to-fine tracking of the zero-crossing curves.

The scale space consists of curves that start at \( \sigma = 0 \) and extend to a certain \( \sigma_{max} \) (infinite \( \sigma_{max} \) is possible) or, conversely, start at a certain \( \sigma_{max} \) and extend to \( \sigma = 0 \). This means that for every edge in the signal a maximum scale can be found that can be considered as a measure of its global importance.

We shall now construct the modified scale space. We begin with the

\[
\text{\( k^{(0)} \) edges at scale } \sigma = 0 : x^{(0)}_1, \ldots, x^{(0)}_k
\]

\( (k^{(0)} \) and \( x^{(0)}_j \) as in the scale space) and we specify for every edge \( x^{(0)}_j \) a
\( \sigma_{\text{max}}(x_j^{(0)}) \), which is the scale at which the zero-crossing curve corresponding to \( x_j^{(0)} \) vanishes (see fig.4.3.1). At each scale \( \sigma \) the modified scale space consists of the edges \( x_j^{(0)} \) for which \( \sigma_{\text{max}}(x_j^{(0)}) \geq \sigma \).

Let these edges be denoted by

\[ \tilde{x}_1^{(\sigma)}, ..., \tilde{x}_n^{(\sigma)} \text{ where } n = k^{(\sigma)} \] (4.3.1)

A segment is defined as

\[ \tilde{I}_j^{(\sigma)} = [\tilde{x}_j^{(\sigma)}, \tilde{x}_{j+1}^{(\sigma)}] \]

The modified scale space (fig.4.3.2) gives an interesting segmentation of the gamma-ray log. As the scale becomes finer new segments appear only within segments that already exist and no new segments appear as the scale becomes coarser. The modified scale space is a multiscale segmentation of the well log. It can be calculated directly from the data and does not depend on a priori knowledge of the signal or on some kind of stochastic model for the signal, in contrast to maximum likelihood methods (Hawkins and ten Krooden[17]).

4.3.3 Multiscale blocking

The next step in the segmentation problem is the assignment of attributes to the segments. A well-known technique in well log analysis is blocking: the replacement of the log values within segments by a constant value (for example the average, the minimum or the maximum value) thus constructing homogeneous layers (Kerzner[18]).

With the modified scale space, blocking at different scales can be achieved as follows. The value assigned to each segment is the average

\[ f_j^{(\sigma)} = \frac{1}{\tilde{x}_{j+1}^{(\sigma)} - \tilde{x}_j^{(\sigma)}} \int_{\tilde{x}_j^{(\sigma)}}^{\tilde{x}_{j+1}^{(\sigma)}} f(x) \, dx , \] (4.3.2)
where the $z_{j+1}^{(\sigma)}$ have been defined in eq. (4.3.1). In fig. (4.3.3) a gamma-ray log is depicted with its modified scale space. The example in fig. 4.3.4 shows that on a coarse scale homogeneous segments can be found, which, on a finer scale, are divided in smaller homogeneous segments.

### 4.3.4 Multiscale trend detection

The loss of trends in the log is a disadvantage of blocking. For that reason it is preferable to make a linear fit within each of the segments. We apply this to the gamma-ray log and its modified scale space depicted in fig. 4.3.5. In each segment the linear least-square-error fit is calculated by

$$\min_{a,b} \int \frac{z_{j+1}^{(\sigma)} - (ax + b)^2}{z_{j}^{(\sigma)}} dx .$$

The result for three values of $\sigma$ is displayed in fig. 4.3.6. For both fine scale and coarse scale, trends are preserved well. The presentation of the log in this way clearly shows the super-position of fine-scale trends on coarse-scale trends.

Of course one could fit other functions to the data in the segments. Also, one could assign additional attributes to the segments, for example the serrateness of the log, measured as the least-square-error of the fit, as the frequency content of the signal, or as the number of fine-scale edges in the segment.

### 4.3.5 Local descent to finer scales

Instead of analyzing the log from its entire (modified) scale space, we propose the method of local descent: we start at a coarse scale in the modified scale space and descend to finer scales only where a stopping criterion is not met.

A possible criterion is a minimum admissible layer thickness. Another possibility works as follows. Within the segments, at a certain scale in the modified scale space, a fit is made. If, in a segment, the error exceeds a threshold, then local descent takes place to a finer scale at which the segment is split up into smaller segments. In this way we aim at an equidistribution of the error. However, there is no guarantee that this goal will be actually achieved. It is possible that in the final result segments exist where the error is still larger than the threshold. Here we will give a demonstration of the local descent method.
using the second criterion in combination with a restricted number of scales. Fig.4.3.7 presents a demonstration of the local descent method in which a blocking is made in the segments. The method uses only the blocking results at $\sigma = 32, 16, 8$ of fig.4.3.4, which means that a minimum admissible scale is present as well. The error in a segment is defined by

$$E_j = \left\{ \frac{1}{\tilde{x}_{j+1}^{(\sigma)} - \tilde{x}_j^{(\sigma)}} \int_{\tilde{x}_j^{(\sigma)}}^{\tilde{x}_{j+1}^{(\sigma)}} (f(x) - \tilde{f}_j^{(\sigma)})^2 \, dx \right\}^{\frac{1}{2}},$$

where $\tilde{x}_j^{(\sigma)}$, $\tilde{x}_{j+1}^{(\sigma)}$ are the edges of the segment and $\tilde{f}_j^{(\sigma)}$ is the mean value of the segment as defined in eq.(4.3.2).

Comparison of fig.4.3.4 and fig.4.3.7 shows that local descent to the finest scale has taken place in the segment extending from approximately 10 to 90 ft. in the blocking result at $\sigma = 32$. In the segment extending from approximately 130 to 200 ft. one step from $\sigma = 32$ to $\sigma = 16$ appears to have been enough. In the remaining segments no local descent was necessary.

4.3.6 Comparison with a common method

In well log analysis a common method for edge detection is to smooth and subsequently differentiate the log curve. An edge is declared where the derivative of the smoothed curve exceeds a predetermined threshold $\alpha$. Some precaution has to be taken to ensure that neighboring points that belong to a part of the log with a large derivative, e.g. a steep trend, are not all declared as edges. Thus, the derivative should also have an extremum in the edge point.

For the derivative of the smoothed log we can write

$$\frac{d}{dx} M(f)(x, \sigma_1) = \frac{1}{\sigma_1^2} \int_{-\infty}^{\infty} f(t) g^{(1)} \left( \frac{x-t}{\sigma_1} \right) dt \quad (4.3.3)$$

where $\sigma_1$ is fixed.

An edge is declared if

$$\left| \frac{d}{dx} M(f)(x, \sigma_1) \right| > \alpha, \quad \alpha > 0, \quad (4.3.4)$$
\[
\frac{d^2}{dx^2} M(f)(x, \sigma_1) = 0,
\]

and

\[
\frac{d^3}{dx^3} M(f)(x, \sigma_1) \neq 0.
\]

The disadvantage of this approach will be demonstrated for the synthetic log of fig.4.3.8, which is presented together with its modified scale space. In the middle of the log there is a segment with a small number of frequency components and a large amplitude. The modified scale space clearly reveals this segment, albeit at a coarse scale. When the common method is applied to this type of signal there are three cases possible, depending on the values of \( \sigma_1 \) and \( \alpha \) in eq.(4.3.3) and eq.(4.3.4).

1. If we think of this log as being a small part of a larger log, then we have the possibility that the smoothing is too strong and the edges that lead to the identification of the middle segment have disappeared.

2. The smoothing corresponds to the scale at which the middle segment is identified in the modified scale space. Then, if the common method is to identify this segment, the absolute value of the derivatives in the two extrema have to be larger than the threshold. If this is the case, the edges will still have a localization error due to the smoothing.

3. The smoothing level corresponds to a scale where too many large absolute values of the derivative are present. This means that a large threshold is required if only a small number of edges are to be found. However, if at an edge the derivative is large, there is no guarantee that this edge is a coarse-scale edge.

In fig.4.3.9 we give a demonstration of the third case. Before differentiation the log has been smoothed with \( \sigma_1 = 2 \). The threshold is such that three edges are detected. A linear fit has been made between the edges. If we compare this result with the result of a linear fit between the edges that are present in the modified scale space at \( \sigma = 64 \) (fig.4.3.10), we conclude that the common method is inferior.
4.4 Conclusions

A method for multiscale segmentation of well logs has been introduced, which fits the multiscale character of well logs, unlike conventional methods.

The method is based on a multiscale transform of the well log, which is a description at multiple levels of smoothness. The multiscale segmentation is achieved by subdividing the log at all scales according to the edges that are associated with the zero crossings of the second derivative. The exact edge locations, which have been lost because of the smoothing, are recovered by coarse-to-fine tracking. Under fairly general conditions the Gaussian appeared to be the only filter in the multiscale transform that does not allow the creation of zero crossings as scale increases. This guarantees that a multiscale segmentation is obtained in which a fine scale segment is part of only one coarse scale segment.

Since in practice a signal is only given at a number of data points we have suggested to write the signal as sum of Gaussians. Imposing an interpolation condition leads to a class of interpolation functions. From this class we select the interpolation function that has the smallest effective width in the frequency domain.

The multiscale segmentation method allows zooming in on parts of the log depending on descent criteria. This yields a segmented log consisting of coarse- and fine-scale segments.

The multiscale segmentation method provides the possibility of performing well log analysis from a new point of view.
Figures section 4.3

Figure 4.3.1: A gamma-ray log (top) with its scale space (bottom). Scale axis is logarithmic and continuous.

Figure 4.3.2: A gamma-ray log (top) with its modified scale space which is a multiscale segmentation (bottom). Scale axis is logarithmic and continuous.
Figure 4.3.3: A gamma-ray log (top) with its modified scale space at scale levels $\sigma = 1, 2, 4, \ldots, 32$.

Figure 4.3.4: Blocking result using three scale levels in the modified scale space: $\sigma = 32$ (top), $\sigma = 16$ (middle), $\sigma = 8$ (bottom). Original gamma-ray log (dotted) is plotted together with results (solid).
Figure 4.3.5: A gamma-ray log (top) with its modified scale space at scale levels $\sigma = 4, 8, 16, ..., 128$.

Figure 4.3.6: Result of linear fit between the edges at 3 scale levels in the modified scale space: $\sigma = 64$ (top), $\sigma = 32$ (middle), $\sigma = 16$ (bottom). Original gamma-ray log (dotted) is plotted together with results (solid).
Figure 4.3.7: Local descent to finer scale using blocking result at $\sigma = 8, 16, 32$ of a gamma-ray log (see fig.4.3.4). Threshold LSQ-error $= 7.8$. Original gamma-ray log (dotted) is plotted together with result (solid).

Figure 4.3.8: A synthetic gamma-ray log (top) with its modified scale-space (bottom).
Figure 4.3.9: Result of linear fit between edges, which have been detected by thresholding the first derivative of a smoothed version of a gamma-ray log. Original gamma-ray log (dotted) is plotted together with result (solid).

Figure 4.3.10: Result of linear fit between edges at scale level $\sigma = 64$ in the modified scale space. Original gamma-ray log (dotted) is plotted together with result (solid).
Chapter 5

Signal matching by dynamic programming

5.1 Introduction

In this chapter the problem is considered of matching signals that are deformed with respect to each other; where the deformation consists of shifts, stretches and compressions. The matching involves the minimization of a certain distance measure or error, between the signals, with respect to the deformation function.

Many minimization algorithms, that are found in literature on signal matching, are based on some kind of gradient method (Martinson et al.[24]). This has the disadvantage that it requires an appropriate initial solution to ensure convergence to the global minimum solution. The minimization algorithm, applied for signal matching in this chapter, is based on dynamic programming (Bellman and Dreyfus[3]). Its application in signal matching has been considered in Leany and Ulrych[21] and Kruse[20]. This method does not suffer from the disadvantages previously mentioned.

The distance measure that is used is important and a number of possibilities are given attention. The distance measure determines what kind of relation between the signals is assumed. For certain distance measures this relation may involve more than only deformation but also certain additive and multiplicative variation. The latter possibility is important especially when signals are to be matched that represent different physical quantities.

The matching (or correlation) of well logs is the area of application that is given special attention. A well log is a sequence of measurements of a certain physical quantity in a bore hole. Examples of physical quantities are the sound velocity the natural gamma radiation and the spontaneous potential. As part of a geophysical survey, well logs are taken from different bore holes in a certain area. Dissimilar rates of sedimentation at the various locations will cause the
well logs to be stretched or compressed with respect to each other. Erosion or faulting may cause a part of a well log to be truncated with respect to the other well logs. In order to obtain a geophysical cross section of the subsurface the deformation function, that connects the well logs, has to be determined.

The organization of this chapter is as follows. In Section 5.2 we begin with a mathematical introduction to the problem. In Section 5.3 a number of distance measures with their properties are given attention. The formulation as a dynamic programming problem is subject of Section 5.4. This involves some approximations that are considered in Section 5.5. A large number of experiments on synthetic data are performed in Section 5.6. In Section 5.7 we consider the matching of well logs.

5.2 Preliminaries

5.2.1 Problem definition

Given two bounded and piecewise continuous functions

\[ f : [0, X] \rightarrow \mathbb{R} , \]
\[ g : [0, Y] \rightarrow \mathbb{R} , \]

the object is to find a match, in some sense, between deformed \( f \) and \( g \).

By deformed \( f \) and \( g \) we mean

\[ f(p(s)) \text{, respectively } g(q(s)) , \]

where \( p \) and \( q \) are continuous functions,

\[ p : [0, S] \rightarrow [0, X] , \]
\[ q : [0, S] \rightarrow [0, Y] . \]

Both \( p \) and \( q \) have to satisfy a condition of causality: they should be monotonically increasing. Furthermore, \( p \) and \( q \) should not both be constant in an interval in \([0, S] \).

If \( D(f, g; p, q) \geq 0 \) is a distance measure between \( f(p) \) and \( g(q) \) then the problem, we are concerned with, is in formula

\[ \min_{p,q} D(f, g; p, q) . \] (5.2.1)
5.2.2 Parametrization-independent measure

Define the distance measure between \( f \) and \( g \) as

\[
D(f, g; p, q) = \int_0^S E(f, g; p, q; s) w(p, q; s) \, ds .
\]

(5.2.2)

Here, \( w \) is a positive weight-function. \( E \) will be called the local distance. Particular choices for \( E \) are considered in Section 5.3. We require that neither \( E \) nor \( D \) depends on the parameterization of the deformation curve. As for now, it is supposed that this is indeed the case for \( E \); here \( D \) is considered.

Suppose that \( r \) is a strictly increasing differentiable function.

\[
r : [0, S] \to [0, R]
\]

Let

\[
\tilde{p}(r(s)) = p(s) ,
\]

\[
\tilde{q}(r(s)) = q(s) ,
\]

\[
\tilde{s} = r(s) , s = r^{-1}(\tilde{s}) ,
\]

\[
d\tilde{s} = r'(r^{-1}(\tilde{s})) \, ds ,
\]

and insert these expressions in eq.(5.2.2). This results in

\[
D(f, g; \tilde{p}, \tilde{q}) = \int_0^R E(f, g; \tilde{p}, \tilde{q}; \tilde{s}) \frac{\tilde{w}(\tilde{p}, \tilde{q}; r^{-1}(\tilde{s}))}{r'(r^{-1}(\tilde{s}))} \, d\tilde{s} .
\]

(5.2.3)

Comparison of eq.(5.2.2) and eq.(5.2.3) shows that \( w \) should be such that

\[
\frac{\tilde{w}(\tilde{p}, \tilde{q}; r^{-1}(\tilde{s}))}{r'(r^{-1}(\tilde{s}))} = w(\tilde{p}, \tilde{q}; \tilde{s}) .
\]

Under the assumption that \( p \) and \( q \) are differentiable a.e., a possibility for \( w \) is

\[
w(p, q; s) = \sqrt{p'(s)^2 + q'(s)^2} .
\]

(5.2.4)

Verification:

\[
\frac{\tilde{w}(\tilde{p}, \tilde{q}; r^{-1}(\tilde{s}))}{r'(r^{-1}(\tilde{s}))} = \frac{\left( \frac{dp(\tilde{s})}{dr^{-1}(\tilde{s})} \right)^2 + \left( \frac{dq(\tilde{s})}{dr^{-1}(\tilde{s})} \right)^2}{r'(r^{-1}(\tilde{s}))} \frac{1}{2}
\]

\[
= \frac{\sqrt{p'(\tilde{s})^2 + q'(\tilde{s})^2}}{r'(r^{-1}(\tilde{s}))} r'(r^{-1}(\tilde{s}))
\]

\[
= w(\tilde{p}, \tilde{q}; \tilde{s}) .
\]
The expression
\[ \int_a^b \sqrt{p'(s)^2 + q'(s)^2} \, ds \]
is the length of the deformation curve \( \Gamma(s) \) for \( s \) between \( a \) and \( b \).

### 5.2.3 Parametrization

The curve that shows the relation between \( f \) and \( g \) is denoted by
\[ \Gamma(s) = \begin{pmatrix} p(s) \\ q(s) \end{pmatrix}. \]
\( \Gamma \) will be called the deformation curve (fig.5.2.1).

![Figure 5.2.1: Deformation curve \( \Gamma \).](image)

Note that in general \( \Gamma \) can not be written as a continuous function of \( x = p(s) \), e.g.
\[ \Gamma(x) = \begin{pmatrix} x \\ v(x) \end{pmatrix}, \]
because \( \Gamma \) vertical now corresponds to a jump-discontinuity in \( v \).
The parameterization that will be used is given by

\[
\begin{align*}
    p(s) &= s - m(s) , \\
    q(s) &= s + m(s) .
\end{align*}
\] (5.2.5)

(See fig.5.2.2.)

![Figure 5.2.2: Choice of parameterization.](image)

When this parameterization is used the distance measure will be denoted by \( D(f, g; m) \) and the local distance measure will be denoted by \( E(f, g; s, m(s)) \). The function \( m \) will be called the deformation function. The causality condition on \( p \) and \( q \) is equivalent with

\[
\frac{|m(s + \Delta s) - m(s)|}{\Delta s} \leq 1 .
\] (5.2.6)

and obviously if \( m \) is differentiable in \( s \):

\[
|m'(s)| \leq 1 .
\] (5.2.7)

### 5.2.4 Uniqueness

The question arises if the solution to the minimization problem in eq.(5.2.1) is unique. In this chapter we will consider distance measures for which, in the most general case,

\[
D(f, g; p, q) = 0 \Leftrightarrow f(p(s)) = C_1 g(q(s)) + C_2 , \quad C_1 \neq 0 .
\]
What becomes clear immediately is that if both $f$ and $g$ are constant then there are infinitely many solutions to the problem. This case of non-uniqueness can be extended. Suppose that $(p(s_1), q(s_1))$ and $(p(s_2), q(s_2))$, with $p(s_2) > p(s_1)$ and $q(s_2) > q(s_1)$, are points on the deformation curve. If $f$ is constant in $(p(s_1), p(s_2))$ and, consequently, $g$ is constant in $(q(s_1), q(s_2))$, then for $s_1 \leq s \leq s_2$, every curve connecting the points $(p(s_1), q(s_1))$ and $(p(s_2), q(s_2))$ is a correct solution. Uniqueness of the solution can be proved, provided these possibilities are excluded.

Henceforth, the parameterization, introduced in the previous section, will be used. We suppose that we have two different solutions to the minimization problem.

\[
f_1(s - m_1(s)) = C_1 g_1(s + m_1(s)) + C_2, \quad C_1 \neq 0, \tag{5.2.8}
\]

\[
f_1(s - m_2(s)) = C_3 g_1(s + m_2(s)) + C_4, \quad C_3 \neq 0, \tag{5.2.9}
\]

with

\[m_1(0) = m_2(0) = 0.\]

Since constant parts have been excluded, eqs. (5.2.8) and (5.2.9) can be transformed into

\[f(s) = g(s),\]

\[f(s - m(s)) = B_1 g(s + m(s)) + B_2, \quad B_1 \neq 0, \text{ with } m(0) = 0.\]

This means that we can consider the possibility that there exists another solution, besides $m(s) = 0$, to

\[g(s - m(s)) = B_1 g(s + m(s)) + B_2, \quad B_1 \neq 0, \text{ with } m(0) = 0.\]

In this case constant parts of $g$ have to be excluded for reasons mentioned before. If $B_1 = 1$ then $B_2 = 0$, otherwise the substitution $g = g_1 + B_2/(1 - B_1)$ can be made. Either way, we obtain an equation of the form:

\[g(s - m(s)) = B g(s + m(s)), \quad B > 0, \text{ with } m(0) = 0. \tag{5.2.10}\]

The case $B < 0$ can be reduced to this equation by considering $\tilde{g} = -g$.

An example of a function $g$ with a non-unique solution $m$ to this equation can easily be found.
If

\[ g(s) = as \text{ with } a \neq 0 \]

then both \( m(s) = 0 \) and

\[ m(s) = \frac{1 - B}{1 + B} s \]

are a solutions of eq.(5.2.10).

It will be proved that if \( m \) has more than one zero then there is only one solution to eq.(5.2.10). In the proof constant parts of \( g \) are excluded.

We begin with analyzing eq.(5.2.10) in the neighborhood of the origin. It is supposed that

\[ \exists \delta_1 > 0 \text{ such that } m(s) \neq 0 \text{ for } s \in (0, \delta_1) . \]

This is not a restriction because there should always exist a point where \( m(s) \neq 0 \) and relocation of the origin will return the former situation. Without loss of generality we take \( m(s) > 0 \) for \( s \in (0, \delta_1) \). Recall that \( m \) is a continuous function. Define the sequence, \( \{s_k\}_{k \in \mathbb{N}} \), by

\[ s_k + m(s_k) = s_{k-1} - m(s_{k-1}) < s_{k-1} , k = 1,\ldots . \]

The begin point is taken arbitrarily but such that

\[ s_0 - m(s_0) \in (0, \delta_1) \]

and

\[ g(s_0 - m(s_0)) \neq 0 . \quad (5.2.11) \]

Such a point can always be found since \( g \) is not constant. The iteration, as \( k \) increases, is depicted in fig.5.2.3. It is clear that

\[ s_k \to 0 \text{ and } m(s_k) \to 0 \text{ as } k \to \infty . \]

For this sequence we find that

\[ g(s_k - m(s_k)) = B g(s_{k-1} - m(s_{k-1})) \]

or

\[ g(s_k - m(s_k)) = B^k g(s_0 - m(s_0)) . \]
Because $g(s_0 - m(s_0)) \neq 0$ by eq. (5.2.11), this sequence only converges as $k$ tends to infinity if $B \leq 1$. If $B = 1$, e.g., if $g(0) \neq 0$, we have that

$$g(0) = g(s_0 - m(s_0))$$.

Because $s_0 - m(s_0)$ is an arbitrary point in $(0, \delta_1)$ this implies that $g$ is constant in that interval. This is in contradiction with the presuppositions.

The case $B < 1$, which implies that $g(0) = 0$, requires more attention. We are left with the problem

$$g(s - m(s)) = B g(s + m(s)) \text{ for } s \in [0, \delta_1),$$

with

$$B < 1, \; g(0) = 0 \text{ and } m(s) > 0 \text{ for } s \in (0, \delta_1).$$

An additional constraint on $m$ has to be imposed:

- $m$ is supposed to have another zero, besides the one in $s = 0$.

The first zero after $s = 0$ is denoted by $s_a$. Hence

$$m(s) > 0 \text{ for } 0 < s < s_a.$$
and
\[ g(s_a) = B g(s_a). \]

Define the sequence
\[ s_k - m(s_k) = s_{k-1} + m(s_{k-1}) > s_{k-1}, \quad k = 1, \ldots \]
The begin point is taken arbitrarily but such that
\[ s_0 + m(s_0) \in (0, s_a) \]
and
\[ g(s_0 + m(s_0)) \neq 0. \]
We have that
\[ s_k \to s_a \text{ and } m(s_k) \to 0 \text{ as } k \to \infty. \]
For this sequence we find that
\[ g(s_k + m(s_k)) = B^{-1} g(s_{k-1} + m(s_{k-1})) \]
or
\[ g(s_k + m(s_k)) = B^{-k} g(s_0 + m(s_0)). \]
Consequently, \( g(s_k + m(s_k)) \) tends to infinity as \( k \to \infty \), which in contradiction with the boundedness of \( g \).

Summary
The uniqueness problem can be addressed using eq.(5.2.10), once constant parts in \( f \) and \( g \) are excluded. It has been proved that there exists only one solution \( m \) to eq.(5.2.10) that has more than one zero. This solution is given by \( m(s) = 0 \).
5.3 Distance measures

5.3.1 General

With $w$ as in eq.(5.2.4), the distance measure $D$ of eq.(5.2.2) is given by

$$D(f, g; p, q) = \int_0^S E(f, g; p, q; s) \sqrt{p'(s)^2 + q'(s)^2} \, ds.$$  \hspace{1cm} (5.3.1)

For the particular choice of parameterization in eq.(5.2.5) this expression becomes

$$D(f, g; m) = \int_0^S E(f, g; s, m(s)) \sqrt{1 + m'(s)^2} \, ds.$$  \hspace{1cm} (5.3.2)

The local distance $E$ determines the sense in which the deformed $f$ and $g$ are to be matched. It determines the relation between the deformed $f$ and $g$ that is to be approximated. Several possibilities will be considered.

5.3.2 Measure for equality

The relation, we are interested in, is

$$f(p(s)) = g(q(s)).$$  \hspace{1cm} (5.3.3)

It is natural to use

$$E(f, g; p, q; s) = |f(p(s)) - g(q(s))|^k,$$  \hspace{1cm} with $k \geq 1.$ \hspace{1cm} (5.3.4)

As required $E$ is parameterization-independent. $D^{1/k}$ is the $L^k$-error between the deformed $f$ and $g$ with a weight function. For $k = 2$, $D$ will be referred to as the square error. The following properties hold for $D$.

1. $D(f, g; p, q) \geq 0$

2. $D(f, g; p, q) = 0$ if and only if $f(p(s)) = g(q(s))$.

The value of $k$ determines the weight that is attached to large local errors, i.e. deviations from the ideal situation in eq.(5.3.3), with respect to small local errors.
Sensitivity

The sensitivity of the deformation function $m$ for distortions using $D$ with $k = 2$ will be examined. We suppose that there exists a distortion of $f$ and $g$ that causes an error in the recovered deformation function. An estimate can be made of the deviation of the recovered deformation function from the original deformation function. We will use the parameterization introduced in eq.(5.2.5). The recovered deformation function is denoted by $\hat{m}$. By assumption we have

$$D(f, g; \hat{m}) < D(f, g; m).$$

Using eq.(5.2.7) we can write

$$\left( \int_0^S [f(s - \hat{m}(s)) - g(s + \hat{m}(s))]^2 ds \right)^{\frac{1}{2}} <$$

$$\left( \sqrt{2} \int_0^S [f(s - m(s)) - g(s + m(s))]^2 ds \right)^{\frac{1}{2}}. \quad (5.3.5)$$

By the mean value theorem

$$f(s - \hat{m}(s)) = f(s - m(s)) - (\hat{m}(s) - m(s)) f'(\xi_{s,m,\hat{m}})$$

and

$$g(s + \hat{m}(s)) = g(s + m(s)) + (\hat{m}(s) - m(s)) g'(\zeta_{s,m,\hat{m}}).$$

By substitution of the previous two equations, the left-hand side of eq.(5.3.5) becomes

$$\left( \int_0^S \left[ \phi \right. \right.$$

$$\left. f(s - m(s)) - g(s + m(s)) - \right.$$
We adopt a shorter notation for this expression: \( ||\phi - \psi||_2 \).

The inequality in eq.(5.3.5) now can be written as

\[
||\phi - \psi||_2 \leq 2^{1\frac{1}{4}} ||\phi||_2 .
\]

Since

\[
||\psi||_2 - ||\phi||_2 \leq ||\phi - \psi||_2 ,
\]

we obtain

\[
||\psi||_2 \leq (1 + 2^{1\frac{1}{4}}) ||\phi||_2 .
\]

Returning to the original notation and squaring both sides of the inequality gives

\[
\begin{align*}
\int_0^S (\hat{m}(s) - m(s))^2 \left[ f'(\xi_{s,m,\hat{m}}) + g'(\zeta_{s,m,\hat{m}}) \right]^2 ds < \\
(1 + 2^{1\frac{1}{4}})^2 \int_0^S [f(s - m(s)) - g(s + m(s))]^2 ds . \tag{5.3.6}
\end{align*}
\]

If now

\[
0 < M \leq [f'(\xi_{s,m,\hat{m}}) + g'(\zeta_{s,m,\hat{m}})]^2 ,
\]

then we arrive at the following inequality

\[
\begin{align*}
\int_0^S (\hat{m}(s) - m(s))^2 ds < \\
\frac{1}{M} (1 + 2^{1\frac{1}{4}})^2 \int_0^S [f(s - m(s)) - g(s + m(s))]^2 ds .
\end{align*}
\]

These inequalities indicate that the deviation of the recovered deformation function from the original deformation function is proportional to the distortion of \( f \) and \( g \), except in intervals where both \( f \) and \( g \) are constant.
5.3.3 Measure for equality with additive variation

Instead of the strong relation between \( f \) and \( g \) in eq. (5.3.3), it might be useful to allow some additive variation. Suppose that \( f \) and \( g \) satisfy the relation

\[
  f(p(s)) = g(q(s)) + \beta(s) .
\]  

(5.3.7)

For certain kind of additive variation \( \beta \) we may still want to find \( p \) and \( q \), or a close approximation, as solution to the minimization problem. The local distance \( E \) determines what kind of behavior of \( \beta \) is considered to be acceptable.

For the local distance \( E \) we propose

\[
  E(f,g;p,q;s) = \\
  \frac{1}{2d} \int_{s_1(s)}^{s_2(s)} |f(p(s)) - g(q(s)) - \{f(p(t)) - g(q(t))\}|^k \sqrt{p'(t)^2 + q'(t)^2} \, dt,
\]  

with \( k \geq 1 \).

(5.3.8)

\( s_1(s) \) and \( s_2(s) \) are defined such that the length of the path used in the calculation before, respectively after, the point \( (p(s),q(s)) \) is constant, that is

\[
  \begin{cases} 
    \int_{s_1(s)}^{s} \sqrt{p'(t)^2 + q'(t)^2} \, dt = d > 0 , \\
    \int_{s_2(s)}^{s} \sqrt{p'(t)^2 + q'(t)^2} \, dt = d > 0 .
  \end{cases}
\]  

(5.3.9)

The local distance \( E \) in eq. (5.3.8) is indeed parameterization-independent. Note that, because we should have \( s_1(s) \geq 0 \) and \( s_2(s) \leq S \), eq. (5.3.9) implies a proper adjustment of the limits of integration in eq. (5.3.1). These new limits of integration will be denoted by \( a \) and \( b \). \( D \) does not measure the deviation of the deformed \( f \) from the deformed \( g \) in just one point but it also compares it to the deviation in an environment of the point \( s: (s_1(s), s_2(s)) \). The parameter \( d \) is free to choose and controls the sensitivity of the distance measure \( D \), as will be discussed further on. For \( k = 2 \), \( D \) will be referred to as the local square error.

\( E \) and \( D \) have the properties:

1. \( E(f,g;p,q;s) \geq 0 \) and \( D(f,g;p,q) \geq 0 \).
2. \( D(f,g;p,q) = 0 \iff E(f,g;p,q;s) = 0 \), \( \forall s \in [a,b] \).

This is equivalent with

\[
f(p(s)) - g(q(s)) = f(p(t)) - g(q(t)) \quad \forall t \in (s_1(s), s_2(s)) \quad \forall s \in (a,b).\]

Subsequently, the 'ideal' relation between \( f \) and \( g \) is

\[
f(p(s)) = g(q(s)) + C, \quad \forall s \in (0,S),
\]

where \( C \) is an arbitrary constant.

**Sensitivity**

In order to examine the sensitivity of \( D \) to additive variation we might be inclined to use the same approach as in the previous section. However, this does not lead to manageable results. Instead we use a simpler approach.

Eq.(5.3.7) is substituted in eq.(5.3.8), after which we obtain for \( D \):

\[
D(f,g;p,q) = \int_a^b \left[ \frac{1}{2d} \int_{s_1(s)}^{s_2(s)} |\beta(s) - \beta(t)|^k \sqrt{p'(t)^2 + q'(t)^2} \, dt \right] \sqrt{p'(s)^2 + q'(s)^2} \, ds. \tag{5.3.10}
\]

If the former expression is small, we expect to find a solution to the minimization problem that is close to the solution \((p,q)\) to the case without additive variation. It can be seen that \( D \) averages local variation, deviation from a constant value, of \( \beta \). The length of every interval \((s_1(s), s_2(s))\) is important for measuring the behavior of \( \beta \). If we recall eqs.(5.2.5), (5.2.7) and (5.3.9) then we find that, for that particular choice of parameterization,

\[
\begin{align*}
  & s_2(s) - s \leq d \leq \sqrt{2} (s_2(s) - s), \\
  & s - s_1(s) \leq d \leq \sqrt{2} (s - s_1(s))
\end{align*}
\]

or

\[
\begin{align*}
  & \frac{d}{\sqrt{2}} \leq s_2(s) - s \leq d, \\
  & \frac{d}{\sqrt{2}} \leq s - s_1(s) \leq d.
\end{align*}
\]
An upper bound for $D$ in eq.(5.3.10) is given by

$$D(f,g;p,q) \leq \frac{1}{d} \int_a^b \int_{s-d}^{s+d} |\beta(s) - \beta(t)|^k \, dt \, ds.$$  

$D$ is small if

$$\frac{1}{d} \int_a^b \int_{s-d}^{s+d} |\beta(s) - \beta(t)|^k \, dt \, ds \ll 1. \quad (5.3.11)$$

This is a weaker statement than eq.(5.3.6) but it does give some indication of the influence of $\beta$. If the value of the integral in eq.(5.3.11) is small we expect that the solution to the minimization problem will be close to the solution with $\beta = 0$. However, in general it is unknown what value of the integral in eq.(5.3.11) can be considered to be small. Experiments may provide more insight.

Eq.(5.3.11) is a condition on the smoothness of $\beta$, a notion which may be clarified by the following. Suppose that $\beta$ is differentiable then linearizing yields

$$\beta(t) = \beta(s) + (t-s)\beta'(s).$$

Substitution of this expression in eq.(5.3.11) results in

$$\frac{2}{k+1} \int_a^b |\beta'(s)|^k \, ds \ll 1. \quad (5.3.12)$$

This makes clear that the larger $d$ is, the smoother $\beta$ should be in order to render a small value of $D$.

### 5.3.4 Measure for equality with multiplicative variation

Allowing multiplicative variation to eq.(5.3.3) gives the expression:

$$f(p(s)) = \alpha(s)g(q(s)). \quad (5.3.13)$$

This time, we want to find $p$ and $q$, or a close approximation, as solution to the minimization problem for certain kind of multiplicative variation $\alpha$. The local distance $E$ determines what kind of behavior of $\alpha$ is considered to be acceptable.
For convenience we define the following quantity:

\[
\langle f, g \rangle = \frac{1}{2d} \int_{s_1(s)}^{s_2(s)} f(t) g(t) \sqrt{p'(t)^2 + q'(t)^2} \, dt , \tag{5.3.14}
\]

where \( s_1 \) and \( s_2 \) are defined as in eq.(5.3.9). The local distance \( E \) that will be used is given by

\[
E(f, g; p, q; s) = 1 - \frac{\langle f(p), g(q) \rangle}{\langle f(p), f(p) \rangle^{1/2} \langle g(q), g(q) \rangle^{1/2}} . \tag{5.3.15}
\]

As required \( E \) is parameterization-independent. The distance measure \( D \) will be referred to as the local correlation measure.

For the properties of \( D \) and \( E \) we find:

1. By Hölder’s inequality,

\[
0 \leq E(f, g; p, q; s) \leq 2 \text{ and }
\]

\[
0 \leq D(f, g; p, q) \leq 2 \int_a^b \sqrt{p'(s)^2 + q'(s)^2} \, ds .
\]

2. \( D(f, g; p, q) = 0 \iff E(f, g; p, q; s) = 0 \forall s \in (a, b) \).

This is the case if and only if

\[
f(p(s)) = C g(q(s)) \forall s \in (0, S) ,
\]

where \( C \) is an arbitrary positive constant.

**Sensitivity**

The sensitivity of \( D \) for certain behavior of \( \alpha \) will be examined in the same way as in Section 5.3.3. After substitution of eq.(5.3.13) in eq.(5.3.15), we arrive at

\[
D(f, g; p, q) = \int_a^b (1 - \frac{\langle \alpha g(q), g(q) \rangle}{\langle \alpha g(q), \alpha g(q) \rangle^{1/2} \langle g(q), g(q) \rangle^{1/2}}) \sqrt{p'(s)^2 + q'(s)^2} \, ds . \tag{5.3.16}
\]
Let us write
\[ \alpha(t) = \alpha(s) + \epsilon(s, t) \]
and
\[ \delta(s) = \sup_{t \in (s_1(s), s_2(s))} |\epsilon(s, t)|. \] (5.3.17)

Substitution of eq.(5.3.17) in eq.(5.3.16) yields
\[
D(f, g; p, q) = \int_a^b \left[ 1 - \frac{1 + \frac{\langle \epsilon g(q), g(q) \rangle}{\alpha(s) \langle g(q), g(q) \rangle}}{\left( 1 + 2 \frac{\langle \epsilon g(q), g(q) \rangle}{\alpha(s) \langle g(q), g(q) \rangle} + \frac{\langle \epsilon g(q), \epsilon g(q) \rangle}{\alpha(s)^2 \langle g(q), g(q) \rangle} \right)^{1/2} \right] \sqrt{p'(s)^2 + q'(s)^2} \, ds. \] (5.3.18)

From eq.(5.3.17) we can derive that
\[
\left| \frac{\langle \epsilon g(q), g(q) \rangle}{\alpha(s) \langle g(q), g(q) \rangle} \right| \leq \frac{\delta(s)}{\alpha(s)},
\]
\[
\left| \frac{\langle \epsilon g, \epsilon g \rangle}{\alpha(s)^2 \langle g(q), g(q) \rangle} \right| \leq \left( \frac{\delta(s)}{\alpha(s)} \right)^2.
\]

Taylor expansion of the integrand in eq.(5.3.18) around
\[ x = \frac{\langle \epsilon g(q), g(q) \rangle}{\alpha(s) \langle g(q), g(q) \rangle} = 0 \quad \text{and} \quad y = \frac{\langle \epsilon g, \epsilon g \rangle}{\alpha(s)^2 \langle g(q), g(q) \rangle} = 0 \]
leads to
\[
D(f, g; p, q) = \frac{1}{2} \int_a^b \left[ \left( \frac{\langle \epsilon g(q), g(q) \rangle}{\alpha(s) \langle g(q), g(q) \rangle} \right)^2 - \frac{\langle \epsilon g, \epsilon g \rangle}{\alpha(s)^2 \langle g(q), g(q) \rangle} + O\left( \left( \frac{\delta(s)}{\alpha(s)} \right)^3 \right) \right] \sqrt{p'(s)^2 + q'(s)^2} \, ds
\]
\[ = O\left( \int_a^b \left( \frac{\delta(s)}{\alpha(s)} \right)^2 \, ds \right). \]

So, it is clear that \( D \) is small if

\[ \int_a^b \left( \frac{\delta(s)}{\alpha(s)} \right)^2 \, ds \ll 1. \tag{5.3.19} \]

This is a condition on the smoothness of \( \alpha \), a notion which may be clarified by the following. Suppose that \( \alpha \) is differentiable then linearizing yields

\[ \alpha(t) = \alpha(s) + (t - s)\alpha'(s) \]

and, by eq.(5.3.17),

\[ \delta(s) = \sup_{t \in (s_1(s), s_2(s))} |(t - s)\alpha'(s)| \]
\[ \leq d |\alpha'(s)|. \]

Eq.(5.3.19) now becomes

\[ d^2 \int_a^b \left( \frac{\alpha'(s)}{\alpha(s)} \right)^2 \, ds \ll 1. \]

The left-hand side of this expression measures the magnitude of the derivative of \( \alpha \) relative to \( \alpha \). It is an increasing function of \( d \), which is the parameter that controls what roughness of \( \alpha \) is admissible.

**Remark**

Some equivalence between the presence of additive and multiplicative variation can be discovered. Substitution of \( \alpha(s) = 1 + \tilde{\beta}(s) \) in eq.(5.3.13),

\[ f(p(s)) = \alpha(s)g(q(s)), \]

results in

\[ f(p(s)) = g(q(s)) + \tilde{\beta}(s)g(q(s)) = g(q(s)) + \beta(s). \]

This means that if the multiplicative variation does not differ much from 1, the local square error as well as the local correlation measure can be used in the matching. A similar assertion can be made for the converse situation.
5.3.5 Measure for equality with additive and multiplicative variation

A combination of the relations between $f$ and $g$ of the previous two sections will be considered, i.e.

$$f(p(s)) = \alpha(s)g(q(s)) + \beta(s) \quad (5.3.20)$$

Define

$$\langle\langle f(p), g(q) \rangle\rangle = \frac{1}{2d} \int_{s_1(s)}^{s_2(s)} \frac{(f(p(s)) - f(p(t)))(g(q(s)) - g(q(t)))\sqrt{p'(t)^2 + q'(t)^2}}{\langle f(p), f(p) \rangle^{\frac{1}{2}} \langle g(q), g(q) \rangle^{\frac{1}{2}}} \, dt. \quad (5.3.21)$$

The local measure $E$ that will be used is

$$E(f, g; p, q; s) = 1 - \frac{\langle\langle f(p), g(q) \rangle\rangle}{\langle f(p), f(p) \rangle^{\frac{1}{2}} \langle g(q), g(q) \rangle^{\frac{1}{2}}} . \quad (5.3.22)$$

Again, $s_1$ and $s_2$ are defined as in eq.(5.3.9). $E$ is parameterization-independent. $D$ will be called the local covariance measure.

For $D$ and $E$ the following properties are important.

1. By Hölder’s inequality,

$$0 \leq E(f, g; p, q; s) \leq 2$$

and

$$0 \leq D(f, g; p, q) \leq 2 \int_{a}^{b} \sqrt{p'(s)^2 + q'(s)^2} \, ds .$$

2. $D(f, g; p, q) = 0 \iff E(f, g; p, q; s) = 0$, $\forall s \in (a, b)$.

This occurs if and only if

$$f(p(s)) - f(p(t)) = C_1(g(q(s)) - g(q(t))), \forall t \in (s_1(s), s_2(s)), \forall s \in (a, b) ,$$

where $C_1$ is an arbitrary positive constant.
This is equivalent with
\[ f(p(s)) = C_1 g(q(s)) + C_2 \quad \forall s \in (0, S), \]
where \( C_1 \) and \( C_2 \) are arbitrary constants but \( C_1 > 0 \).

**Sensitivity**

The sensitivity of \( D \) for small deviations from an ideal case will be examined. We introduce
\[ \delta(s) = \sup_{t \in (s_1(s), s_2(s))} |\alpha(t) - \alpha(s)| \]
and
\[ \eta(s) = \sup_{t \in (s_1(s), s_2(s))} |\beta(t) - \beta(s)|. \]

A derivation analogous to the one in the previous section involving substitution of eq.(5.3.20) in eq.(5.3.22) and Taylor expansion leads to
\[ D(f, g; p, q) = O\left( \int_a^b \left( \frac{\eta(s)}{\alpha(s)} \right)^2 + \left( \frac{\delta(s)}{\alpha(s)} \right)^2 \, ds \right). \]

The derivation of this equation is omitted. \( D \) is small if
\[ \int_a^b \left( \frac{\eta(s)}{\alpha(s)} \right)^2 + \left( \frac{\delta(s)}{\alpha(s)} \right)^2 \, ds \ll 1. \quad (5.3.23) \]

An approximation of this expression is
\[ d^2 \int_a^b \left( \frac{\beta'(s)}{\alpha(s)} \right)^2 + \left( \frac{\alpha'(s)}{\alpha(s)} \right)^2 \, ds \ll 1. \]
5.4 Dynamic Programming

5.4.1 General

The deformation curve $\Gamma$ is the minimum-distance path in the rectangle with corner points $(0,0),(X,0),(0,Y)$ and $(X,Y)$, connecting $(0,0)$ and $(X,Y)$. (See fig. 5.4.1.) The minimum-distance path will be determined using Dynamic Programming (DP). DP is based on Bellman’s principle of optimality (Bellman and Dreyfus [3]). For our problem this principle can be formulated as: "If $B$ is on the optimal path from $A$ to $C$ then the part of that path between $B$ and $C$ is the optimal path from $B$ to $C$, independent of how one arrives in $B."$ The elaboration of this principle is the subject of the next section.

![Diagram showing deformation curve $\Gamma$ with points $A$, $B$, $C$ and paths $(0,0)$, $(X,0)$, $(0,Y)$, $(X,Y)$]

Figure 5.4.1: The optimal path from $B$ to $C$ is part of the optimal path from $A$ to $C$.

In the remaining part of this chapter only the parameterization of eq.(5.2.5) will be used:

\[
\begin{align*}
p(s) &= s - m(s), \\
q(s) &= s + m(s).
\end{align*}
\]
5.4.2 Formulation as DP-problem

Suppose that we have an approximation of eq.(5.3.2) of the form

\[
\hat{D}(f,g; \hat{m}) = \sum_{k=0}^{N-1} E(f,g; s_k, \hat{m}(s_k)) l(\hat{m}(s_{k+1}) - \hat{m}(s_k)),
\]

where \( s_k \) and \( \hat{m} \) are given by

\[
s_k = k \Delta s + a \ (b = a + N \Delta s),
\]

\[
\hat{m}(s_k) = \text{entier} \left[ \frac{m(s_k)}{\Delta s} + \frac{1}{2} \right] \Delta s.
\]

From eq.(5.2.6) we derive that only three (i.e. \(-\Delta s, 0, \Delta s\)) values are allowed for \( \hat{m}(s_{k+1}) - \hat{m}(s_k) \). The function \( l \) accounts for the weight-function in \( D \). The justification of the approximations, as well as the function \( l \), is subject of Section 5.5.

The approximating minimization problem can be formulated as

\[
\min_{\hat{m}(s_0), \ldots, \hat{m}(s_N)} \sum_{k=0}^{N-1} E(f,g; s_k, \hat{m}(s_k)) l(\hat{m}(s_{k+1}) - \hat{m}(s_k))
\]

subject to \( \hat{m}(s_{k+1}) - \hat{m}(s_k) = -\Delta s, 0, \Delta s \ k = 0, \ldots, N - 1 \)

Next we define

\[
\Phi(\hat{m}(s_N), N) = \min_{\hat{m}(s_0), \ldots, \hat{m}(s_{N-1})} \sum_{k=0}^{N-1} E(f,g; s_k, \hat{m}(s_k)) l(\hat{m}(s_{k+1}) - \hat{m}(s_k))
\]

subject to \( \hat{m}(s_{k+1}) - \hat{m}(s_k) = -\Delta s, 0, \Delta s \ k = 0, \ldots, N - 1 \).

Eq.(5.4.3) now can be formulated as

\[
\min_{\hat{m}(s_N)} \Phi(\hat{m}(s_N), N)
\]

For \( \Phi \) it is possible to write

\[
\Phi(\hat{m}(s_N), N) = \min_{\hat{m}(s_{N-1})} \{ \Phi(\hat{m}(s_{N-1}), N - 1) + E(f,g; s_{N-1}, \hat{m}(s_{N-1})) l(\hat{m}(s_N) - \hat{m}(s_{N-1})) \}
\]

subject to \( \hat{m}(s_N) - \hat{m}(s_{N-1}) = -\Delta s, 0, \Delta s \).
A general recursion-formula is
\[ \Phi(\hat{m}(s_n), n) = \min_{\hat{m}(s_{n-1})} \{ \Phi(\hat{m}(s_{n-1}), n - 1) + E(f, g; s_{n-1}, \hat{m}(s_{n-1})) l(\hat{m}(s_n) - \hat{m}(s_{n-1})) \} \]
subject to \( \hat{m}(s_n) - \hat{m}(s_{n-1}) = -\Delta s, 0, \Delta s \) \hspace{1cm} (5.4.4)
\[ \Phi(\hat{m}(s_0), 0) \overset{\text{def}}{=} 0 \].

This recursion is important because it states that the overall minimization problem of eq.(5.4.3) can be reduced to \( N \) minimizations involving only one variable.

Consider fig.5.4.2, where a part of the rectangle is depicted with grid points for \( k = 0, \ldots, K \). In order to determine the original number of search possibilities we define \( b_k \) as the number of possible paths from the bottom left corner to the grid points corresponding to \( k \). It is straightforward to observe that \( b_k \) satisfies the recurrence relation,
\[ b_{k+1} = 3b_k \ , \quad k = 1, \ldots, K - 1 \text{ with } b_1 = 3 \ , \]
and consequently
\[ b_k = 3^k \quad \text{for } k = 1, \ldots, K \ . \]

With the recursion-formula of eq.(5.4.4) the number of searches has decreased considerably.

Denote the number of searches necessary in this case to connect the points at \( k \) to the bottom left corner by \( c_k \). Suppose we want to connect the paths at \( k \) to \( k+1 \). Because of eq.(5.4.4) the number of searches \( c_{k+1} \) is equal to the number of searches thus far, \( c_k \), plus the number of searches to connect \( k \) to \( k+1 \), which is just the number of grid points at \( k \) times 3. In formula
\[ c_{k+1} = c_k + 3k \ , \quad k = 1, \ldots, K - 1 \text{ with } c_1 = 3 \ . \]

The solution of this recurrence relation is given by
\[ c_k = \frac{3}{2}k(k - 1) + 3 \quad \text{for } k = 1, \ldots, K \ . \]

So we observe that, whereas \( b_k \) has exponential dependence on \( k \), \( c_k \) has only quadratic dependence.

What is missing in eq.(5.4.4) is the fact that in our problem the path is restricted to a rectangle, this can be incorporated by adding some constraints.
5.4.3 Remarks on minimization

A different approach to the minimization problem can be found in Martinson et al.[24]. The deformation function is represented by a sum of basis functions multiplied by coefficients. Say

\[ m = \sum_{i=1}^{N} a_i \phi_i \]

The coefficients are determined by a search algorithm, which is based on the necessary condition for the solution that the gradient with respect to those coefficients is zero. In formula,

\[ \min_{a_1, \ldots, a_N} D(f, g; a_1, \ldots, a_N) . \] (5.4.5)

A necessary condition is expressed by

\[ \frac{\partial D}{\partial a_i} (f, g; a_1, \ldots, a_N) = 0 \quad \text{for } i = 1, \ldots, N . \] (5.4.6)

In general the object function \( D(f, g; a_1, \ldots, a_N) \) will have several local minima for which the previous equation is satisfied by a certain set of parameters. Evaluation of all local minima in order to obtain the global minimum is a tedious task, which can be avoided by supplying an initial guess of the correct solution.
to the minimization algorithm. From the initial guess the algorithm converges
to the correct solution, the global minimum. We note that in Martinson et
al.\cite{martinson2000} the causality condition $|m'| \leq 1$ is not present. Incorporation would
further complicate the minimization problem.

The advantage of the formulation of the minimization problem as a DP-problem
is that none of the previously mentioned problems are present. The recursion
formula in eq.(5.4.4) solves the minimization problem exactly; it always renders
the global-minimum solution.

However, a price has been paid to accomplish this. In eqs.(5.4.5) and (5.4.6)
there are no restrictions on the parameterization of the deformation function. If
the minimization problem is to be formulated as a DP-problem, a certain known
ordering of the parameters is required. The ordering enables to minimize for
one parameter at the time and to use the result in the subsequent minimization.
In the case of Section 5.4.2 we have made use of the causality condition on
the deformation curve, which results in a temporal ordering of the parameters
$m(s_k)$, $k = 0, \ldots, N$.

5.5 Discretization

5.5.1 The integral expression

The object of this section is to arrive at eq.(5.4.1) by discretization of eq.(5.3.2).
This equation can be written as

$$D(f, g; m) = \sum_{k=0}^{N-1} \int_{s_k}^{s_{k+1}} E(f, g; s, m(s)) \sqrt{1 + m'(s)^2} \, ds . \tag{5.5.1}$$

Here, $s_k$ has been defined in eq.(5.4.2). If we assume that $E(f, g; s, m(s))$ is
differentiable with respect to $s$ on $(s_k, s_{k+1})$ then, by the mean value theorem,

$$E(f, g; s, m(s)) = E(f, g; s_k, m(s_k)) + \mu_1(s)(s - s_k) ,$$

with $s \in [s_k, s_{k+1})$ and where

$$|\mu_1(s)| \leq \max_{[s_k, s_{k+1})} \left| \frac{d}{ds} E(f, g; s, m(s)) \right| = M_{1,k} .$$

Eq.(5.5.1) now becomes
\[ D(f, g; m) = \sum_{k=0}^{N-1} \left[ \int_{s_k}^{s_{k+1}} \sqrt{1 + m'(s)^2} \, ds + e_{1,k} \right]. \]

\( e_{1,k} \) satisfies

\[
e_{1,k} = \int_{s_k}^{s_{k+1}} \mu_1(s) (s - s_k) \sqrt{1 + m'(s)^2} \, ds \leq \frac{M_{1,k}}{\sqrt{2}} \Delta s^2. \tag{5.5.2}
\]

It is also assumed that \( E(f, g; s, m(s)) \) is continuously differentiable with respect to \( m \) in every interval \([\hat{m} - \frac{1}{2} \Delta s, \hat{m} + \frac{1}{2} \Delta s]\), where \( \hat{m} \) is defined in eq.(5.4.2). In that case

\[ E(f, g; s_k, m(s_k)) = E(f, g; s_k, \hat{m}(s_k)) + \mu_2(m) (m(s_k) - \hat{m}(s_k)) \]

with \( \frac{1}{2} \Delta s \leq \hat{m} - m < \frac{1}{2} \Delta s \) and where

\[ |\mu_2(m)| \leq \sup_{\frac{1}{2} \Delta s \leq \hat{m} - m < \frac{1}{2} \Delta s} \left| \frac{d}{dm} E(f, g; s_k, m(s_k)) \right| = M_{2,k}. \]

For eq.(5.5.1) we finally obtain

\[ D(f, g; m) = \sum_{k=0}^{N-1} \left[ E(f, g; s_k, \hat{m}(s_k)) \int_{s_k}^{s_{k+1}} \sqrt{1 + m'(s)^2} \, ds + e_{1,k} + e_{2,k} \right]. \tag{5.5.3}
\]

Here \( e_{1,k} \) satisfies eq.(5.5.2) and \( e_{2,k} \) satisfies

\[
e_{2,k} = (m(s_k) - \hat{m}(s_k)) \mu_2(m) \int_{s_k}^{s_{k+1}} \sqrt{1 + m'(s)^2} \, ds \leq \frac{M_{2,k}}{\sqrt{2}} \Delta s^2.
\]

Most formulas for \( E \), that we have considered, are extensive and do not provide very manageable expressions for \( M_1 \) and \( M_2 \), in order to estimate the \( e_{1,k} \) and \( e_{2,k} \) in eq.(5.5.3). It is possible to cope with a simple possibility, the square error,
\[ E(f, g; s, m(s)) = |f(s - m(s)) - g(s + m(s))|^2. \]

Then
\[ \left| \frac{d}{ds} E(f, g; s, m(s)) \right| = 2 \left| f(s - m(s)) - g(s + m(s)) \right|. \]

\[ |f'(s - m(s))(1 - m'(s)) - g'(s + m(s))(1 + m'(s))| \]

and
\[ \left| \frac{d}{dm} E(f, g; s, m(s)) \right| = \]
\[ 2|f(s - m(s)) - g(s + m(s))||f'(s - m(s)) + g'(s + m(s))|. \]

These equations show that the relative magnitude of \( e_{1,k} \) and \( e_{2,k} \) is proportional to the derivative of \( f \) and \( g \) and that the relative magnitude of \( e_{1,k} \) is also proportional to the derivative of \( m \). For the other choices of \( E \) we restrict ourselves to the assumption that the observation we have made concerning the influence of the derivatives of \( f, g \) and \( m \) also applies to the other possibilities for \( E \).

### 5.5.2 The weight function

The last step from eq.(5.3.2) to its approximation in eq.(5.4.1) involves a discussion on the term:

\[ l = \int_{s_k}^{s_{k+1}} \sqrt{1 + m'(s)^2} \, ds. \]

It is a problem to give a good estimate of the length of \( m \) in the interval \((s_k, s_{k+1})\) since only \( \hat{m}(s_k) \) and \( \hat{m}(s_{k+1}) \) are known. Consider the situation on the grid at a point \((s_k, \hat{m}(s_k))\) in fig.5.5.1. There are two essentially different possibilities. They are given in fig.5.5.1 and denoted by \( l_1 \) and \( l_2 \). For \( l_1 \) and \( l_2 \) the areas are depicted where the curves lie that give rise to either connection. If the curve is approximated as a straight line in that interval, the most obvious length estimate would be the value \( \Delta s \) for \( l_1 \) and \( \Delta s \sqrt{2} \) for \( l_2 \). However, it can be shown that for the set of straight lines resulting in \( l_1 \) the estimate of \( \Delta s \) is structurally to small and that for \( l_2 \) the estimate \( \Delta s \sqrt{2} \) is structurally to large. Information concerning this problem can be obtained in Dorst and Smeulders [11], in which an overview is given of length estimators for straight
lines represented on a (square) grid, with some remarks on generalization to other curves. The essence of the underlying theory is the following.

Through each area, corresponding either to \( l_1 \) or to \( l_2 \) lines are drawn at random. The expectation of the length of these lines is used as the estimate. Of course one has to specify what random lines are. Here, we refer to Duda and Hart [12]. The estimate of the length of a line segment obtained in this way is given by 1.06 for \( l_1 \) and 1.18 for \( l_2 \), values which have first been derived in Groen and Verbeek [15]. In this way, however, each line segment is considered as being independent of the rest of the curve. Especially for straight lines this approach is too rude. It has to be admitted that experiments did not show major differences in the matching results for the various length estimates. In the final matching algorithm the formula for \( l \) in eq.(5.4.1) we will use is

\[
l(x) = \begin{cases} 
1.06 & \text{if } x = 0, \\
1.18 & \text{if } |x| = \Delta s. 
\end{cases}
\]

5.5.3 Problems for some distance measures

Up to now it did not matter, in the steps that have been made to discretize eq.(5.3.2), whether we were dealing with a forward problem (deformation curve known) or with an inverse problem (deformation curve to be determined). However, in the case of an inverse problem the calculation of \( E(f, g; s_k, \hat{m}(s_k)) \)
raises problems if it is one of the possibilities in eqs. (5.3.8), (5.3.15) or (5.3.22), because they all involve the calculation of a certain mean value over a part of the deformation curve; a curve that is unknown. In order to show how this problem is dealt with a particular local distance will be considered.

From eq. (5.3.8), with \( k = 2 \) and with \( p \) and \( q \) as in eq. (5.2.5), we obtain

\[
E(f, g; s_k, \hat{m}(s_k)) =
\]

\[
\frac{1}{2d} \int_{s_1(s_k)}^{s_2(s_k)} |f(s_k - \hat{m}(s_k)) - g(s_k + \hat{m}(s_k)) - \{f(t - m(t)) - g(t + m(t))\}|^2 \sqrt{1 + m'(t)^2} \, dt,
\]

with

\[
\begin{cases}
  s_k^d \\
  s_1(s_k) \\
  s_2(s_k) \\
  s_k
\end{cases}
\int_{s_1(s_k)}^{s_2(s_k)} \sqrt{1 + m'(t)^2} \, dt = d.
\]

(5.5.4)

(5.5.5)

An approximation of \( m(t) \) in the interval \( (s_1(s_k), s_2(s_k)) \) has to be made that enables the evaluation of eqs. (5.5.4) and (5.5.5). The only reasonable option is to approximate \( m(t) \) by a constant value in the interval \( (s_1(s_k), s_2(s_k)) \),

\[
m(t) \approx \hat{m}(s_k),
\]

\[
m'(t) \approx 0 \text{ for } t \in (s_1(s_k), s_2(s_k)).
\]

This leads to

\[
\begin{cases}
  s_1(s_k) = s_k - d, \\
  s_2(s_k) = s_k + d
\end{cases}
\]

and

\[
E(f, g; s_k, \hat{m}(s_k)) \approx \frac{1}{2d} \int_{s_k-d}^{s_k+d} |f(s_k - \hat{m}(s_k)) - g(s_k + \hat{m}(s_k)) - \{f(t - \hat{m}(s_k)) - g(t + \hat{m}(s_k))\}|^2 \, dt.
\]

(5.5.6)
The accuracy of the approximations in the previous formulas is inverse proportional to $\Delta s$, $m'$, $f'$, $g'$ and $d$. The consequence of the approximations is that the local distance in eq.(5.5.6) is calculated at a point in the grid, $(s_k, \hat{m}(s_k))$, using not-deformed $f$ and $g$.

To check on the effect of this approximation and possibly to improve on the result an iterative procedure can be used:

- The deformation function that has been calculated is used, possibly smoothed, to deform $f$ and $g$.

- For the deformed $f$ and $g$ a new deformation curve is calculated.

- This procedure is repeated until the deformation becomes negligible.

Smoothing of the deformation function can be done by averaging,

$$m_{\text{smooth}}(s) = \frac{1}{2l} \int_{-l}^{l} m(s + \tau) \, d\tau$$

(5.5.7)

Because the averaging crosses the boundaries of the interval $(0, S)$, $m$ is extended beyond these boundaries,

$$m(s) = m(0) \text{ for } s \leq 0 ,$$

$$m(s) = m(S) \text{ for } s \geq S .$$

With $\tau_n = n\Delta s$ and $l = \hat{l}\Delta s$ and using eq.(5.4.2), we obtain the following approximation

$$\hat{m}_{\text{smooth}}(s_k) = \frac{1}{2\hat{l} + 1} \sum_{n=-\hat{l}}^{\hat{l}} \hat{m}(s_k + \tau_n) .$$

(5.5.8)

The integral in eq.(5.5.6) has to be discretized as well. It will be done in the same way with the same stepsize $(\Delta s)$ as before. It naturally follows that $d$ should be a multiple of $\Delta s$.

It will be found in the experiments that the smoothing has a positive effect on the results. In the experiments the iterative procedure did not appear to be necessary.

As we have noted, a large value for $d$ should be avoided because the local
distances can only be calculated using not-deformed signals. However, small values for \( d \) should be avoided as well. In general, we expect that a (local) distance measure that is insensitive to certain variations cannot distinguish between signal and variation if the window within which the calculations take place is too small. As the measure is insensitive to more kinds of variations the window length should be larger.

With respect to the discretization, we can observe this problem in more detail. Consider eq.(5.5.6) and suppose that \( d \) is small such that both \( f \) and \( g \) can be approximated by a straight line within the interval \((s_k - d, s_k + d)\). Substitution of this approximation in eq.(5.5.6) leads to disappearance of the deformation function \( \hat{m}(s_k) \) from the expression. Exactly the same effect occurs in the local covariance measure as can be seen by inspection of eq.(5.3.21). However, it is not present in the local correlation measure.

5.6 Experiments on synthetic data

5.6.1 Preliminary

Construction of data

In the experiments two functions, denoted by \( f \) and \( g \), will be matched. The function \( f \) is constructed by linear interpolation between a number of data points \( f_k = f(k), k = 0, \ldots, K \),

\[
f(x) = \sum_{k=0}^{K} f_k \, l_k(x) \quad \text{with } x \in [0, K].
\] (5.6.1)

The linear interpolation functions, \( l_k \), are defined by

\[
l_k(x) = \begin{cases} 
  x - (k - 1) & \text{for } k - 1 < x \leq k, \\
  (k + 1) - x & \text{for } k < x < k + 1, \\
  0 & \text{for } k = 1, \ldots, K - 1 \text{, elsewhere},
\end{cases}
\]

\[
l_0(x) = \begin{cases} 
  1 - x & \text{for } 0 \leq x < 0, \\
  0 & \text{elsewhere}
\end{cases}
\]

and

\[
l_K(x) = \begin{cases} 
  x - (K - 1) & \text{for } K - 1 < x \leq K, \\
  0 & \text{elsewhere}.
\end{cases}
\]
The function $g$ is constructed by

$$g(x) = f(v(x)) = \sum_{k=0}^{K} f_k l_k(v(x)) \quad \text{with} \ x \in [0, K].$$

Both $f$ and $g$ will be sampled at points:

$$x_n = \frac{n}{N} K \quad \text{with} \ n = 0, \ldots, N.$$ 

With reference to eq.(5.4.2), we note that the discretization width is given by

$$\Delta s = \frac{K}{N}. \quad (5.6.2)$$

In the forthcoming experiments two different set of data points $\{f_k\}_{k=0, \ldots, K}$ will be used. For both of them, $K = 128$ in eq.(5.6.1).

**Variables**

The influence on the matching result of the following quantities is subject of investigation in the experiments:

- The discretization width: $\Delta s = \frac{K}{N}$.

- The derivative of the deformation function: $m'$.

- The window length in the local distance measure: $2d$.

- The averaging length used for smoothing the recovered deformation function: $2l$.

In Sections 5.6.4, 5.6.5, 5.6.6, 5.6.7 and 5.6.8 a number of quantities will be added to this list.

**Error**

The following quantities can be calculated to evaluate the matching result.

- Error in the calculation of $m$:

$$E_m = \frac{1}{K} \int_0^K (m(s) - \hat{m}(s))^2 \, ds.$$
In which $m$ denotes the original or model deformation function and in which $\hat{m}$ denotes the calculated or recovered deformation function. For other variables the same convention is adopted.

- When more convenient the error in $v$:

$$E_v = \frac{1}{K} \int_0^K (v(x) - \hat{v}(x))^2 \, dx.$$  

- With $\hat{g}(x) = f(\hat{v}(x))$:

$$E_g = \frac{1}{K} \int_0^K (\hat{g}(x) - g(x))^2 \, dx.$$  

5.6.2 Experiment 1

Deformation function

The deformation function $v$ that is used in this section is:

$$v(x) = \begin{cases} 
  x & \text{for } 0 \leq x \leq \frac{1}{20} K , \\
  a(x - \frac{1}{20} K) + \frac{1}{20} K & \text{for } \frac{1}{20} K < x \leq x_{\text{int}} , \\
  \frac{1}{a}(x - \frac{19}{20} K) + \frac{19}{20} K & \text{for } x_{\text{int}} < x < \frac{19}{20} K , \\
  x & \text{for } \frac{19}{20} K \leq x \leq K , 
\end{cases}$$  \hspace{1cm} (5.6.3)

with $a > 0$ and where $x_{\text{int}}$ is given by

$$x_{\text{int}} = \frac{K}{20}(a - \frac{19}{a} + 19)/(a - 1/a).$$

From $v$ we can derive that the $m$, for which $f(s - m(s)) = g(s + m(s))$, is given by

$$m(s) = \begin{cases} 
  0 & \text{for } 0 \leq s \leq \frac{1}{20} K , \\
  \alpha(s - \frac{1}{20} K) & \text{for } \frac{1}{20} K < s \leq \frac{1}{2} K , \\
  -\alpha(s - \frac{19}{20} K) & \text{for } \frac{1}{2} K < s < \frac{19}{20} K , \\
  0 & \text{for } \frac{19}{20} K \leq s \leq K , 
\end{cases}$$  \hspace{1cm} (5.6.4)
where
\[ \alpha = \frac{1 - a}{1 + a}. \]

Because of eq.(5.2.7), we should have \(-1 \leq \alpha \leq 1\). If \(\alpha\) is negative the first part of \(f\) is compressed and the second part of \(f\) is stretched to make \(g\). If \(\alpha\) is positive it is the other way around.

Results

Example 1

In the first example \(N = 256\) and \(a = 1.5\). The applied distance measure is the square error. Consider fig.5.6.1. The function \(f\) is depicted at the bottom of the picture with the \(x\)-axis horizontally and the \(y\)-axis pointing upwards. The function \(g\) is depicted at the left side of the picture with the \(x\)-axis vertically and the \(y\)-axis pointing to the right.

The recovered deformation curve is shown as a solid curve; the model deformation function is shown as a dotted curve. It can be seen that the recovered deformation curve matches the model deformation curve very well.

The deviation of \(\hat{m}\) from \(m\) is shown in fig.5.6.2. We have also mapped the model deformation function onto the grid, using eq.(5.4.2). The result is denoted by \(\tilde{m}\). The deviation of \(\tilde{m}\) from \(m\) is shown in fig.5.6.3. It can be seen in fig.5.6.2 that in most points \(m - \hat{m}\) is of the same magnitude as \(m - \tilde{m}\). The behavior of \(m - \tilde{m}\) is such that smoothing \(\hat{m}\) will improve the result.

With the recovered deformation function, \(\hat{g}\) has been determined. In advance, the deformation curve has been smoothed with window length \(2l = 6\Delta s\). \(f\) and \(g\) are shown in fig.5.6.4. \(\hat{g}\) together with \(g\) is depicted in fig.5.6.5. We observe that a very good match has been achieved. The important data of this test are condensed in the following table.

<table>
<thead>
<tr>
<th></th>
<th>square error</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(N)</td>
</tr>
<tr>
<td></td>
<td>256</td>
</tr>
<tr>
<td></td>
<td>smoothing</td>
</tr>
<tr>
<td></td>
<td>with</td>
</tr>
<tr>
<td></td>
<td>without</td>
</tr>
<tr>
<td></td>
<td>(E_m)</td>
</tr>
<tr>
<td></td>
<td>0.005</td>
</tr>
</tbody>
</table>
For $\tilde{m}$ the quantity $E_m$ equals 0.018.

Example 2

The second example is given in fig.5.6.6. The deformation curve is steeper than in the previous example. $\tilde{m}$ deviates more from $m$ than in the previous example; this occurs especially in the part where $f$ has been compressed. The explanation is that a proper representation of a steeper deformation curve requires a finer grid. In addition, we have that the part of $g$ that consists of a compressed part of $f$ is quite rough and therefore requires a finer discretization as well. The functions $f$ and $g$ are shown in fig.5.6.7. $g$ and $\hat{g}$ are displayed in fig.5.6.8. We conclude that $\hat{g}$ matches $g$ very well indeed. The important data of this test are condensed in the following table.

<table>
<thead>
<tr>
<th>square error</th>
</tr>
</thead>
<tbody>
<tr>
<td>$N$</td>
</tr>
<tr>
<td>256</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>smoothing</th>
</tr>
</thead>
<tbody>
<tr>
<td>with</td>
</tr>
<tr>
<td>without</td>
</tr>
</tbody>
</table>

| $E_m$ | $E_g$ | $E_m$ | $E_g$ |
| 0.008 | $2.0 \times 10^{-4}$ | 0.032 | $5.8 \times 10^{-4}$ |

Example 3

In this example we have applied the local square error. The deformation curve is the same as in Example 1. The result is shown in fig.5.6.9. In this picture the deformation function $\tilde{m}$ has been smoothed. Fig.5.6.10 contains $f$ and $g$. Fig.5.6.11 contains $g$ and $\hat{g}$.

<table>
<thead>
<tr>
<th>local square error</th>
</tr>
</thead>
<tbody>
<tr>
<td>$N$</td>
</tr>
<tr>
<td>512</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>smoothing</th>
</tr>
</thead>
<tbody>
<tr>
<td>with</td>
</tr>
<tr>
<td>without</td>
</tr>
</tbody>
</table>

| $E_m$ | $E_g$ | $E_m$ | $E_g$ |
| 0.013 | $2.7 \times 10^{-4}$ | 0.048 | $1.22 \times 10^{-3}$ |
The results obtained with the square error are better than the results obtained with the local square error, despite the finer discretization in the latter case.

Error versus discretization

For each of the four distance measures (square error, local square error, local correlation and local covariance) we have determined $E_m$ as a function of $N$, which is related to the discretization width by eq.(5.6.2),

$$\Delta s = \frac{K}{N}.$$  

In these tests, $a = 1.5$ and $d = 1$. The result is depicted in fig.5.6.12. The recovered deformation function has not been smoothed. Obviously, the error $E_m$ decreases as $N$ increases. It is remarkable that the local covariance measure renders a smaller error for $N = 256$, $N = 512$ and $N = 1024$ than the local square error and the local correlation measure; this is not coincidental but structural. It will also be observed in the experiments in Section 5.6.3. For $N = 128$, however, the local covariance measure performs considerably worse than the other measures.

The same experiments have been performed with window length $d = 2$. The results are shown in fig.5.6.13. For $N \geq 256$, the square error performs best and the local covariance performs second best. For $N = 128$ it is remarkable that the square error performs worst.

Comparison of the two figures also reveals that the value of the error $E_m$ for the local square error, the local correlation measure and the local covariance remains larger or takes longer to decrease for larger $N$ if $d = 2$ than if $d = 1$.

We conclude that these observations are in accordance with considerations in Section 5.5.3. However, we do not want to attach much weight to the results for $N = 128$.

Smoothing the deformation function

The recovered deformation function is smoothed, by averaging over a certain interval, before it is used to deform $f$ to make $\hat{g}$ (see eq.(5.5.7)). In fig.5.6.14 the errors $E_m$ and $E_g$ as function of $I$, which is half the averaging length, have
been depicted. The general data belonging to the experiments are:

<table>
<thead>
<tr>
<th>local square error</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
</tr>
<tr>
<td>512</td>
</tr>
</tbody>
</table>

It can be seen that as l tends away from 0 there is an increase of accuracy. Apparently, the averaging suppresses small noisy errors in the deformation curve. Beyond a certain value of l the errors begin to increase. The averaging begins to affect the deformation function itself instead of the small deviations.

Error versus window length

The error $E_m$ as a function of half the window length $d$, that is used in the calculation of the local square error, the local correlation measure and the local covariance measure is depicted in fig.5.6.15. Calculations have been done for three values of $a$. In all calculations $N = 512$. The following observations can be made:

1. The error increases as $a$ increases. Hence, as the deformation curve becomes steeper.

2. For a small window length the local covariance measure performs much worse than the other two measures; the local correlation measure performs better than the local square error.

3. For larger window lengths the local covariance measure performs better than the other two measures, especially for steeper deformations.

The first two observations are in accordance with the assertions of Section 5.5.3. The third observation can be explained from the insensitivity of the local covariance measure to both additive and multiplicative variation. This insensitivity may cause the measure to be less sensitive to deviations caused by the discretization than the other local distance measures.
5.6.3 Experiment 2

Deformation function

The second deformation function \( v \) is

\[
v(x) = \frac{b}{2} \left( \frac{x}{K} \right)^2 \left( \frac{x}{K} - 1 \right)^2 (2 \frac{x}{K} - 1) + x , \quad \text{with} \quad x \in [0, K].
\]  

(5.6.5)

This function satisfies

\[
v(0) = 0, \quad v(K) = K, \quad v\left(\frac{1}{2}K\right) = \frac{1}{2}K, \quad v'(0) = 1 \quad \text{and} \quad v'(K) = 1.
\]

It has points of inflexion in

\[
x = \frac{1}{2}K \quad \text{with} \quad v'(x) = 1 + \frac{b}{16},
\]

\[
x = \frac{2}{2\sqrt{15} + 10} \quad \text{with} \quad v'(x) \approx 1 - 0.05b
\]

and for

\[
x = \frac{\sqrt{15}}{10} + \frac{1}{2} \quad \text{with} \quad v'(x) \approx 1 - 0.05b .
\]

The causality condition requires that \( v'(x) \geq 0 \). This condition is met if

\[-16 \leq b \leq 20
\]

Results

Example

The matching result applying the local correlation measure is shown in fig.5.6.16. \( f \) and \( g \) are shown in fig.5.6.17, \( g \) and \( \hat{g} \) are depicted in fig.5.6.18. A good match has been found. The recovered deformation function exhibits a somewhat larger deviation from the model deformation curve near the beginning and near the end. We do not have a proper explanation for this phenomenon.

<table>
<thead>
<tr>
<th>( N )</th>
<th>( \Delta s )</th>
<th>( b )</th>
<th>( d )</th>
</tr>
</thead>
<tbody>
<tr>
<td>512</td>
<td>0.25</td>
<td>10</td>
<td>1.25</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>smoothing</th>
<th>( l )</th>
<th>2.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>with</td>
<td></td>
<td></td>
</tr>
<tr>
<td>without</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( E_v )</td>
<td>( E_g )</td>
<td>( E_v )</td>
</tr>
<tr>
<td>0.037</td>
<td>3.2 ( 10^{-4} )</td>
<td>0.115</td>
</tr>
</tbody>
</table>
Error versus window length

The test results are given in fig.5.6.19 for $b = 10$ and for two values of $N$. Clearly, a larger $N$ renders smaller errors.
The observations that can be made with respect to the influence of $d$ are the same as observations 2 and 3 of the previous section. They therefore affirm the surmises of Section 5.5.3.

Smoothing the deformation function

For this example we have determined the error $E_g$ as a function of the averaging length $2l$. We have applied the local correlation measure with $b = 10$ and $d = 1.25$ for $N = 256$ and $N = 512$. In fig.5.6.20, we observe that the error $E_g$ behaves in a fashion comparable to fig.5.6.14.

5.6.4 Noise

Data

The function $f$ that is used in this section is the same as in Section 5.6.2. The function $g$ is constructed as has been described in Section 5.6.1 using the deformation function in eq.(5.6.3). In this section $f$ and $g$ will be matched after they have been corrupted with additive Gaussian white noise. Hence, we will match $\tilde{f}$ and $\tilde{g}$, that are given by

$$\tilde{f}(x) = f(x) + n_1(x),$$
$$\tilde{g}(x) = g(x) + n_2(x),$$

where $n_1$, as well as $n_2$, is Gaussian white noise with variance $\sigma^2$. The relative magnitude of the noise can be measured by a properly defined noise-to-signal ratio, for example

$$NSR_f = \frac{\sigma}{S_f},$$

in which

$$S_f = \left( \frac{1}{K} \int_0^K (f(x) - \tilde{f})^2 \, dx \right)^{\frac{1}{2}},$$

with

$$\tilde{f} = \frac{1}{K} \int_0^K f(x) \, dx.$$
**Results**

**Example**

A matching result involving two signals, to which a particular realization of noise has been added, will be given. The general data of this experiment are:

<table>
<thead>
<tr>
<th>$N$</th>
<th>$\Delta s$</th>
<th>$a$</th>
<th>$d$</th>
<th>$S_f$</th>
<th>$S_g$</th>
<th>$\sigma$</th>
</tr>
</thead>
<tbody>
<tr>
<td>512</td>
<td>0.25</td>
<td>0.8</td>
<td>1.25</td>
<td>0.297</td>
<td>0.284</td>
<td>0.04</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>smoothing</th>
<th>l</th>
<th>3.75</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>with</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$E_m$</td>
<td>$E_g$</td>
<td>0.520</td>
<td>$7.5 \times 10^{-4}$</td>
<td>$E_m$</td>
<td>$E_g$</td>
<td>$1.5 \times 10^{-3}$</td>
</tr>
</tbody>
</table>

The average value of $E_m$, without smoothing, over a number of realizations of noise with the same variance is 0.400. This is less than in the example. The error $E_g$ has been calculated using the signals without noise. If the noise is retained the error becomes $E_g = 2.6 \times 10^{-3}$ with and without smoothing.

The model and recovered deformation curve, together with the two noisy signals are depicted in fig.5.6.21. We observe a major deviation of the recovered deformation curve from the model deformation curve in the part that corresponds to the segment in between the first two major lumps in $\bar{f}$ and $\bar{g}$. This can be explained by the fact that the noise-to-signal ratio is large in that particular segment. In fig.5.6.22 the two noisy signals are depicted. The recovered smoothed deformation has been used to deform $f$ itself. The result together with $g$ is shown in fig.5.6.23. As was to be expected the fit in the segment in between the first two major lumps is not good.

**Error versus $\sigma$**

A matching has been performed with various realizations of noise, for a number of values of $\sigma$. For every $\sigma$ the square root of the average value of the error $E_m$ has been calculated. In the experiments the same $f$ and deformation function $m$ as in the previous example have been used. In fig.5.6.24 the square root of the average value of the error $E_m$ as function of $\sigma$ is shown for matching in which the square error has been used. In fig.5.6.25 we have done the same for matching in which the local covariance has been applied with $d = 1.25$. 
Especially in fig.5.6.24 we observe a relation between $\sigma$ and $\sqrt{E_m}$ that is approximately linear. This can be understood by examination of eq.(5.3.6), in which we observed that the deviation of the recovered deformation from the model deformation function is proportional to the distortion of $f$ and $g$.

### 5.6.5 Additive variation

**Data**

In this section the function $f$ is the same as in Section 5.6.2. The function $g$ is constructed by

$$g(x) = f(v(x)) + \beta(x) \text{ with } x \in [0, K] ,$$

where the deformation function $v$ is given in eq.(5.6.3). The additive variation $\beta$ we will use is a sine,

$$\beta(x) = B \sin(\omega x) ,$$

(5.6.6)

The number of periods on the interval $[0, K]$ is denoted by $N_c$ and equals $K \omega / 2\pi$.

In Section 5.3.3 we have attempted to gain insight in the influence of additive variation on the matching. Of major importance appeared to be the derivative of the additive variation and the window length of the local measure (eq.(5.3.12)). In this section we will be concerned with the influence of $B$, $N_c$ and $d$ on the matching result. The error $E_g$ is defined in this section as

$$E_g = \frac{1}{K} \int_0^K (\hat{g}(x) - f(v(x)))^2 dx , \text{ with } \hat{g}(x) = f(\hat{v}(x)) .$$

(5.6.7)

**Results**

**Example 1**

The general data of this experiment are:

<table>
<thead>
<tr>
<th>N</th>
<th>$\Delta s$</th>
<th>$a$</th>
<th>$d$</th>
<th>$A$</th>
<th>$B$</th>
<th>$N_c$</th>
</tr>
</thead>
<tbody>
<tr>
<td>512</td>
<td>0.25</td>
<td>1.75</td>
<td>1.25</td>
<td>-</td>
<td>0.2</td>
<td>3</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$E_m$</th>
<th>$E_g$</th>
<th>$E_m'$</th>
<th>$E_g'$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.070</td>
<td>$3.4 \times 10^{-4}$</td>
<td>0.192</td>
<td>$1.95 \times 10^{-2}$</td>
</tr>
</tbody>
</table>
In fig.5.6.26, we show the signals \( f \) and \( g \). It can be seen that the addition of a sine with relatively large amplitude has corrupted the signal \( g \) considerably. Both \( f \) and \( g \) have four major lumps and we can see that especially the third and fourth lump are distorted. Furthermore, we observe that a downward trend is present in the wiggly segment in between the first two lumps of the signal \( g \) which is not present in the corresponding segment in \( f \). In fig.5.6.27 the functions \( m \) and smoothed \( \tilde{m} \) are shown. In fig.5.6.28, we show the signals \( \hat{g} \) and \( g \). From these pictures we conclude that the match is quite good. The major lumps in \( \hat{g} \) have been placed on the corresponding lumps in \( g \) (fig.5.6.28). We note that the wiggly segment has been placed correctly as well. However, as can be seen in fig.5.6.27 the recovered deformation curve deviates considerably from the model deformation curve within this segment. This is reflected in \( \hat{g} \) in fig.5.6.28. The difference \( g(x) - \hat{g}(x) \) is depicted in fig.5.6.29. The result is the original sine \( \beta \) with some distortion.

**Example 2**

A second example example is given in which a sine with a smaller amplitude and a larger frequency has been added to one of the signals.

<table>
<thead>
<tr>
<th>( N )</th>
<th>( \Delta s )</th>
<th>( a )</th>
<th>( d )</th>
</tr>
</thead>
<tbody>
<tr>
<td>512</td>
<td>0.25</td>
<td>1.75</td>
<td>0.75</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>smoothing with</th>
<th>( l )</th>
<th>3.75</th>
</tr>
</thead>
</table>

\[
\begin{array}{c|c|c|c|c|}
 E_m & E_g & E_m & E_g \\
\hline
 0.014 & 4.8 \times 10^{-4} & 0.097 & 1.57 \times 10^{-3} \\
\end{array}
\]

The results are depicted in figs.5.6.30, 5.6.31, 5.6.32 and 5.6.33. In this case we do not observe large deviations of the smoothed \( \tilde{m} \) from \( m \). The results are better than in the previous example. This is entirely due to the smaller amplitude \((B)\) of the added sine.

**Influence of the frequency of the additive variation**

In fig.5.6.34 we have depicted the error \( E_m \) as a function of half the window length \( d \) for a number of values of the number of periods \( N_c \) with \( a = 1.5 \) and
$B = 0.2$. The applied distance measure is the local square error. We will list a number of observations.

1. The larger $N_c$ is the sooner $E_m$ begins to increase.

2. The larger $N_c$ is the faster $E_m$ increases.

In fig.5.6.35 the results are shown for $a = 1.75$ and $B = 0.05$. In this case we observe the following.

1. For $N_c = 0$, $N_c = 1$, $N_c = 5$ the behavior is comparable to the previous case with $a = 1.5$ and $B = 0.2$ (fig.5.6.34).

2. For $N_c = 10$ $E_m$ has an oscillating behavior, still with a relatively small error for $d = 0.75$.

3. For $N_c = 20$ $E_m$ shows behavior deviant from what has been observed thus far. No clear upward trend is present. The error is minimal for a relatively large value of $d$.

From these observations, we draw the following conclusions.

1. If the frequency of the added sine is relatively small, hence if the added variation is smooth, the matching results as a function of $d$ behave as expected:

   (a) The error $E_m$ has a minimum for a certain $d$.

   (b) The location of this minimum tends to shift to smaller $d$ as the frequency of the added variation increases. This indicates when the local square error becomes aware of the added variation.

2. If the frequency of the added variation is large and the amplitude is sufficiently small the matching does not necessarily render a bad result. However, the influence of the window length is different than before. This can be explained by noting that the local square error has been designed to correct for slowly varying additive variations. A high-frequency sine deviates from this model and can better be considered as additive noise.

It has to be stressed that we have to be careful in extending these conclusions to general cases. We have observed in other experiments that they hold in a global sense but that deviations are possible. In general, it is advisable to compare the matching results for different values of $d$. 
5.6.6 Multiplicative variation

Data

In this section the function $f$ is the same as in Section 5.6.2. The function $g$ is constructed by

$$g(x) = \alpha(x)f(v(x)) \text{ with } x \in [0, K],$$

where the deformation function $v$ is given in eq.(5.6.3). The $\alpha$ that is used is given by

$$\alpha(x) = (1 + A \sin(\omega x)) \text{ with } x \in [0, K] \text{ and } -1 < A < 1. \quad (5.6.8)$$

The number of periods on the interval $[0, K]$ is again denoted by $N_c$ and equals $K \omega/2\pi$. The error $E_g$ is defined as in eq.(5.6.7).

Results

Example

The example we give in which one of the signals has been multiplied by a sine is rather extreme but illustrative. The signals that are to be matched are depicted in fig.5.6.36. The distortion in $g$ caused by the multiplicative variation is considerable. The general data of the experiment are:

<table>
<thead>
<tr>
<th>$N$</th>
<th>$\Delta s$</th>
<th>$a$</th>
<th>$d$</th>
<th>$A$</th>
<th>$B$</th>
<th>$N_c$</th>
</tr>
</thead>
<tbody>
<tr>
<td>256</td>
<td>0.5</td>
<td>1.2</td>
<td>6</td>
<td>0.3</td>
<td>-</td>
<td>10</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$E_m$</th>
<th>$E_g$</th>
<th>$E_m$</th>
<th>$E_g$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.039</td>
<td>7.8 $10^{-4}$</td>
<td>0.095</td>
<td>1.1 $10^{-3}$</td>
</tr>
</tbody>
</table>

The smoothed recovered deformation curve and the model deformation curve are depicted in fig.5.6.37. We can see in fig.5.6.37 that $\hat{m}$ matches $m$ very well. The functions $g$ and $\hat{g}$ are depicted in fig.5.6.38. We observe that a good match has been achieved. An interesting observation can be made in the two segments in between the first three major lumps. It can be seen that the corresponding segments in $g$ and $\hat{g}$ do not show much resemblance. However, they have been placed at the correct location. Apparently, the four
major lumps in the signals are enough for the matching algorithm to find a correct match.

**Influence of the frequency of the multiplicative variation**

In fig.5.6.39 we have depicted the error $E_m$ as a function of half the window length $d$ for a number of values of $N_c$ with $a = 1.2$ and $A = 0.3$. The behavior of $E_m$ is similar to its behavior in fig.5.6.34. For observations and conclusions we therefore refer to the previous section.

### 5.6.7 Additive and multiplicative variation

**Data**

In this section the function $f$ is the same as in Section 5.6.2. The function $g$ is constructed by

$$g(x) = \alpha(x)f(v(x)) + \beta(x) \text{ with } x \in [0, K].$$

where the deformation function $v$ is given in eq.(5.6.3), $\alpha$ is given in eq.(5.6.8) and $\beta$ is given in eq.(5.6.6). The error $E_g$ is defined as in eq.(5.6.7).

**Results**

**Example**

The signals $f$ and $g$ are shown in fig.5.6.40. The recovered and model deformation function are shown in fig.5.6.41. $g$ together with $\hat{g}$ are depicted in fig.5.6.42. The general data of this example are given by:

<table>
<thead>
<tr>
<th>$N$</th>
<th>$\Delta s$</th>
<th>$a$</th>
<th>$d$</th>
<th>$A$</th>
<th>$B$</th>
<th>$N_c$</th>
</tr>
</thead>
<tbody>
<tr>
<td>256</td>
<td>0.5</td>
<td>0.8</td>
<td>2.5</td>
<td>0.4</td>
<td>0.2</td>
<td>2.5</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$E_m$</th>
<th>$E_g$</th>
<th>$E_m$</th>
<th>$E_g$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.007</td>
<td>$2.1 \times 10^{-4}$</td>
<td>0.037</td>
<td>$5.4 \times 10^{-4}$</td>
</tr>
</tbody>
</table>

Obviously, the deformation function has been determined correctly. A good match has been achieved between $g$ and $\hat{g}$. 
Influence $N_c$

In fig.5.6.43 we have depicted the error $E_m$ as a function of half the window length $d$ for a number of values of $N_c$ with $a = 0.8$, $A = 0.4$ and $B = 0.2$. Again, we observe a similar behavior as in Section 5.6.5. An initial decrease of the error $E_m$ followed by an increase beyond a certain value of $d$. As $N_c$ increases the break point shifts to smaller $d$ and the increase of the error becomes larger.

5.6.8 Deformation with shift

General

An extreme matching problem occurs if a part of one of the signals is absent in the other signal. This is a limit case in which the deformation function is vertical or horizontal. By a proper parameterization and a corresponding formulation as a DP-problem the matching algorithm is capable of handling horizontal and vertical paths through the error matrix.

A problem arises if we evaluate the error along such a path. This will be exemplified using the figure at the top of the next page. In this figure a deformation curve is depicted which relates a function $f$, to a function $g$, which is thought to be along the $y$-axis. At $y = \frac{1}{2}$, a part of $f$ over a length $L$ is absent in $g$. For that part the deformation function is horizontal. Put in other words, a proper match is achieved by shifting a part of $f$ until it matches the corresponding part of $g$.

The deformation function $m$ is defined, on the interval $[0,1]$, by

$$m(s) = \begin{cases} 
0 & 0 \leq s \leq 0.05, \\
\frac{L}{4} - \frac{L}{4} (s - 0.05) & 0.05 < s < \frac{1}{2} - \frac{L}{4}, \\
\frac{1}{2} - s & \frac{1}{2} - \frac{L}{4} \leq s \leq \frac{1}{2} + \frac{L}{4}, \\
\frac{L}{4} - \frac{L}{4} (s - 0.95) & \frac{1}{2} + \frac{L}{4} < s < 0.95, \\
0 & 0.95 \leq s \leq 1.
\end{cases} \quad (5.6.9)$$

with $0 \leq L \leq 0.9$. 

The functions $f$ and $g$ are related by
\[ f(s - m(s)) = g(s + m(s)) \quad \text{for } 0 \leq s < \frac{1}{2} - \frac{L}{4} \text{ and } \frac{1}{2} + \frac{L}{4} < s \leq 1. \]

The part of $f$ that has no match in $g$ is given by
\[ f(2s - \frac{1}{2}) \quad \text{for } \frac{1}{2} - \frac{L}{4} \leq s \leq \frac{1}{2} + \frac{L}{4}. \]

Calculation of the square error along the entire path results in
\[ D(f, g; m) = \sqrt{2} \int_{\frac{1}{2} - \frac{L}{4}}^{\frac{1}{2} + \frac{L}{4}} (f(2s - \frac{1}{2}) - g(\frac{1}{2}))^2 \, ds = \frac{1}{\sqrt{2}} \int_{\frac{1}{2} - \frac{L}{4}}^{\frac{1}{2} + \frac{L}{4}} (f(x) - g(\frac{1}{2}))^2 \, dx \quad (5.6.10) \]
and this quantity is unequal zero unless
\[ f(s) = g(\frac{1}{2}) \quad \text{for } \frac{1}{2} - \frac{L}{4} \leq s \leq \frac{1}{2} + \frac{L}{4}. \]

Eq.(5.6.10) determines the variation of $f$ around $g(\frac{1}{2})$ within the interval $(\frac{1}{2} - \frac{1}{2}L, \frac{1}{2} + \frac{1}{2}L)$ and is an increasing function of $L$. The matching algorithm will only retrieve the original deformation curve as long as the quantity in eq.(5.6.10) is the minimum value of the error, hence if
\[ D(f,g;m) = \frac{1}{\sqrt{2}} \int_{\frac{1}{2}-\frac{1}{4}}^{\frac{1}{2}+\frac{1}{4}} (f(x) - g(\frac{1}{2}))^2 \, dx < D(f,g;\mu) \quad \forall \mu \neq m \quad (5.6.11) \]

If \( L \) becomes too large this is likely to become untrue. Similar results hold for other distance measures. For a large value of \( L \) we have depicted the (square) error matrix for two signals constructed using the deformation function of eq.(5.6.9) in fig.5.6.44. It is unsatisfactory if the matching algorithm is not able to deal with a long shift, because the error matrix consists of two clearly distinguishable subpaths where the error is almost zero.

In Section 5.3.2 we have noted that the value of \( k \) in eq.(5.3.4) determines the weight that is attached to large local errors. We consequently expect that a lower value of \( k \) in eq.(5.3.4) will give better results in the case of a deformation with a shift. The same can be expected for the local error in eq.(5.3.8) and the local correlation and covariance measures could be adapted in this way as well.

A step towards a more general solution to the problem would be to exchange the global approach, in which the entire signals are matched, for a local approach, in which the signals are matched in parts. This approach induces the problem of the selection of those parts.

We may select pairs of points in the error matrix and determine for each pair the optimal path between the points. In this way we obtain a number of subpaths which can be used for the final matching solution. The problem we encounter is twofold. Firstly, which points do we select, hence which parts of the signal do we match, and secondly, how do we choose from the available subpaths. An approach like this was used in Leany[21].

The selection of parts of the signals that have to be matched may be done using a certain segmentation algorithm (see for example Vermeer and Alkemade[31], [32]). However, in general there is no guarantee that the part of one of the signals that is absent in the other is recognized as, or can even considered to be, a single segment. It has to be stressed that this problem does not originate from the minimization method that is used, i.e. dynamic programming. Another minimization method, as the one in Section 5.4.3, suffers from the same difficulties.

For the local square error, local correlation measure and local covariance measure similar results as in eq.(5.6.11) can be obtained. However, for these
measures another important problem exists. In Section 5.5.3 we have noted
that because of discretization these measures are in fact calculated using
non-deformed signals. For large values of \(|m'|\) this may cause errors. In case of a
horizontal or vertical deformation curve this effect is clearly present. If the local
square error, the local correlation measure or the local covariance measure is
calculated along the model deformation curve the part of \(f\) that is absent in
\(g\) is compared to parts of \(g\) around \(y = \frac{1}{2}\). An iterative procedure as has been
proposed is not the solution to this problem.

Data

The deformation function that will be used in the experiments satisfies

\[ m(s) = K \mu\left(\frac{s}{K}\right) \text{ with } s \in [0, K], \]

where \(\mu\) is the deformation function defined in eq.(5.6.9).

Results

Example 1

In figs.5.6.45, 5.6.46 and 5.6.47 we have depicted an example for which the
general data are:

<table>
<thead>
<tr>
<th>square error</th>
</tr>
</thead>
<tbody>
<tr>
<td>(N)</td>
</tr>
<tr>
<td>512</td>
</tr>
<tr>
<td>smoothing</td>
</tr>
<tr>
<td>with</td>
</tr>
<tr>
<td>(E_m)</td>
</tr>
<tr>
<td>-</td>
</tr>
</tbody>
</table>

In fig.5.6.45 we observe that the deformation function has been recovered very
well. In fig.5.6.46 the functions \(f\) and \(g\) are displayed. We see that the part
of \(f\) approximately in between 58 and 70 is absent in \(g\). In fig.5.6.47 it can be
seen that \(\hat{g}\) matches \(g\). A small peak in \(\hat{g}\) with respect to \(g\) in the vicinity of
65 marks the location onto which the points of \(f\) in between 58 and 70 have
been mapped.
Example 2

In this example the length of the shift has been increased with respect to the length in the previous example. The general data are:

<table>
<thead>
<tr>
<th></th>
<th>square error</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
</tr>
<tr>
<td>512</td>
<td>0.25</td>
</tr>
</tbody>
</table>

The matching result is shown in fig.5.6.48. The recovered deformation function deviates considerably from the model deformation function. In fig.5.6.49 we have depicted the square error along the recovered and model deformation function. In formula,

\[ E_1(f, g; s, m(s)) = (f(s - m(s)) - g(s + m(s)))^2 \]

and

\[ E_2(f, g; s, \hat{m}(s)) = (f(s - \hat{m}(s)) - g(s + \hat{m}(s)))^2. \]

For \( m \) we observe large values of \( E_1 \) in the interval corresponding to the shift and very small values elsewhere. For \( \hat{m} \) we observe very small values of \( E_2 \) for \( 0 < s < 50 \) because in this interval \( \hat{m} \) fits \( m \). For \( s > 50 \) \( E_2 \) is small compared to the peaks in \( E_1 \).

Example 3

With the same shift length as in the previous example, we now apply the error in eq.(5.3.4) with \( k = 1 \).
| \( k = 1 \) error |  
|---|---|---|---|
| \( N \) | \( \Delta s \) | \( L \) | \( d \) |
| 512 | 0.25 | 0.16 | - |
| smoothing | \( l \) | - |
| with | without |
| \( E_m \) | \( E_g \) | \( E_m \) | \( E_g \) |
| - | - | 0.202 | 3 \( \times 10^{-4} \) |

The deformation curve is shown in fig.5.6.50. The result is considerably better than in fig.5.6.48. In fig.5.6.51 we show the error of eq.(5.3.4), with \( k = 1 \), calculated along the recovered and model deformation curve. The result for the model deformation curve is smaller than in fig.5.6.51, which explains why the recovered curve is closer to the model curve than if the square error is applied.

**Example 4**

In fig.5.6.52 an example has been given in which the local covariance measure has been applied. General data:

| local covariance m. |  
|---|---|---|---|
| \( N \) | \( \Delta s \) | \( L \) | \( d \) |
| 512 | 0.25 | 0.08 | 2.5 |
| smoothing | \( l \) | - |
| with | without |
| \( E_m \) | \( E_g \) | \( E_m \) | \( E_g \) |
| - | - | 0.072 | 8.1 \( \times 10^{-4} \) |

For the main part the deformation function has been retrieved well. However, the shift has not been found correctly. In general, matching across a shift, in which the local square error, the local correlation measure or the local covariance measure is used, will not give considerably more problems than matching using the square error. However, the exact location and length of the shift will not be found accurately.

**5.6.9 Final remarks**

The experiments on synthetic data that have been carried out in this section have confirmed the expectations concerning dependence of the results on the
discretization width, small distortions, and window length of the local square error and the local correlation/covariance measure in relation to the smoothness of the additive and/or multiplicative variation. These expectations that were based on the theoretical considerations in Sections 5.3 and 5.5.

The quantification of the relation between additive and multiplicative variation and the window length used in the local square error and the local correlation/covariance measure has not been realized.

In practice the choice of distance measure depends on the expected properties of the signals that have to be matched. Because the local covariance measure is the most generally suited distance measure, this will often be the obvious choice.

5.7 Application to well logs

5.7.1 Introduction

Well logs that have been obtained from different bore holes can exhibit differences, such as stretches, compressions and shifts, due to different geological circumstances at the various locations. In Section 5.7.2 the matching method will be tested on three pairs of well logs. For the study of the results we visually select some major features, such as edges and peaks, in the well logs.

An important problem in well log processing is the segmentation (or zoning) of well logs. It has been suggested to make a segmentation of the well logs first and to use the segmentation for the matching (Shaw and Cubitt[28]). This is in contrast to the pointwise approach we have presented thus far. Matching with prior segmentation can be done by assigning to each segment in one well log a segment in the other well log. If the length of the segments differs either of them can be (linearly) stretched. This is certainly a sound approach, although it induces a number of problems.

The first problem is one of compatibility, which in a simple form arises if for instance a part of one of the well logs has been subdivided in two segments, while the corresponding part in the other well log has been subdivided in three. It is clear that a matching method based on segmentation should be able to assign three segments in one well log to two in another well log, or to adjust the segmentation during the matching procedure.

A second problem is the deformation within segments. The validation of the
match between two segments has to be done by considering all possible deformations that map the mutual begin and edge points onto each other. These problems can be overcome if the length of the segments is small, but this is a return to a pointwise method.

In Section 5.7.3 we propose to use a segmentation of the well logs to evaluate the matching result. Suppose the deformation curve has been calculated and that the well logs have been deformed accordingly. The distance measure as a function of depth can now be calculated using the deformed well logs. As we know, the overall result is optimal, but this does not mean that the value of the distance measure is small everywhere. We want to distinguish between parts of the well logs that match well and parts that do not match well. We are not interested in points or very small intervals, but in significant parts where the mean value of the distance measure is small and in which a few bad outliers can be tolerated. In addition, we would like the subdivision in intervals of good and bad match to fit to a subdivision or segmentation of the well logs themselves. To this purpose we use the multiscale segmentation method for well logs that has been studied in Chapter 4.

5.7.2 Matching of two well logs

Data

Experiments are carried out on three pairs of well logs:

- 2 gamma-ray logs,
- 2 spontaneous potential (SP) logs,
- a spontaneous potential and an acoustic impedance log.

In all pictures the x-axis unit is $\frac{1}{2}$ ft. which is also the sample width.

In the well logs that are displayed at the top we will give markers, such as edges and peaks. The deformation curves determines which points in the well logs correspond to each other. In the figures we will show onto which points in the bottom well log the edges and peaks in the top well log are mapped.

The distance measure

For all the matching experiments we use the local covariance measure, because we do not want the result to be influenced by smooth amplitude variations in
the well logs and because we want to match two well logs of different physical quantities.
The local covariance measure will in general be the measure that is most fit for well log correlation. An exception to this can be the problem of correcting the depth difference between well logs that have been taken with different trains. Although the measured physical quantity and the well are the same, the logs may show differences as a result of dilation of the cable and jamming of the sondes (Vermes[35]). This matching problem is probably best addressed using the square error or the error in eq.(5.3.4) with some other value of $k$.

After calculation of the deformation curve, one of the well logs is deformed. Before and after the deformation we calculate for each point the local covariance and its average value over the entire length of the logs. The local covariance is defined by

$$C(x) = \frac{< f, g >(x)}{< f, f >^{1/2} < g, g >^{1/2}},$$

where

$$< f, g >(x) = \frac{1}{2d} \int_{x-d}^{x+d} (f(t) - f(x))(g(t) - g(x)) \, dt.$$

This is just 1 minus the local covariance measure defined in eq.(5.3.21) for not-deformed signals. The value of the local covariance is bounded,

$$-1 \leq C(x) \leq 1.$$

Its average value is given by

$$\bar{C} = \frac{1}{X} \int_{0}^{X} C(x) \, dx.$$

Two gamma-ray logs

The result for two gamma-ray logs is shown in fig.5.7.1. The window length that has been used in the calculation is 15 ft., $d = 15$. The gamma-ray logs are depicted in the middle and the bottom of the picture. At the top of the picture the local covariance is shown. Markers have been selected for the upper gamma-ray log. Dashed lines point to these markers and to points in the lower
log onto which these markers have been mapped. Corresponding points are connected with solid lines.

It can be observed that most peaks, edges and segments in the top well log have been mapped onto very similar parts of the bottom well log. The local covariance that has been depicted at the top shows a rough spiky behavior with a low average value.

The result after deformation is shown in fig.5.7.2. We can now make a good comparison between the parts of the well log that have been matched. We observe that the match of the markers and their interjacent segments is satisfactory. The awkward part at the end of the bottom well log is due to the stretching of a small segment.

In the newly calculated local covariance we observe a considerable increase of the average value. The local covariance as a function of depth is close to its maximum in a large number of points. The histogram of the local covariance before and after deformation is shown in fig.5.7.7.

It can be seen that although two edges may be correctly mapped onto each other this does not imply that the interjacent segments show a large value of the local covariance. We also observe that the local covariance is still spiky. This is disadvantageous if we want to determine segments with a good or a bad match (See Section 5.7.3).

**Two SP logs**

In the previous section the two gamma-ray logs appeared to have a lot of features and segments in common. In the two SP logs in this section there is far less similarity (see fig.5.7.3). The window length that has been used in the calculation is 15 ft., \( d = 15 \). We observe that the local covariance is again a spiky function with a low average value. Some of the segments that have been mapped onto each other show some similarity on visual inspection. However, others do not show resemblance. It has to be stressed that this does not imply that the matching method does not function well. The method only determines the best possible match given two signals.

The result after deformation is shown in fig.5.7.4. The average value of the local covariance has increased considerably. Some parts of the well logs can
be seen to have a large similarity. The local covariance remains a spiky curve. The histogram of the local covariance before and after deformation is shown in fig.5.7.8.

A SP log and an impedance log

In this section a SP log and an impedance log will be matched. It is rather hazardous to try to match logs that correspond to different physical quantities. However, if similar structures are present in the logs we hope they will be matched correctly.

The window length that is used in the calculation is 10 ft., \( d = 10 \). Consider fig.5.7.5. We have marked the parts of the SP log which have been matched to similar log shapes in the impedance log. For other parts of the logs it is hard or even impossible to observe any similarity. The picture after deformation is clearer (See fig. 5.7.6). More similar log shapes can be distinguished. The spiky local covariance shows some parts of the logs where its value is large. The histogram of the local covariance before and after deformation is shown in fig.5.7.9.

5.7.3 Multiscale evaluation

General

It is important in the evaluation of the matching result to distinguish between parts of the well log that match well and parts that do not match well. The tool for measuring the match is the local covariance calculated after deformation. The detection of parts of good match could be done by determination of parts in the local covariance of high and approximately constant value. The disadvantage of this approach is that the information and structure contained in the well logs is almost completely neglected.

An important aspect of the evaluation is to obtain segments that are of significant length and that a few outliers should not influence the detection of parts of good or bad match.

We propose a method of evaluation that incorporates a segmentation of the well logs. It is based on the modified scale space as discussed in Section 4.3.2 and the method of local descend originating from it (Section 4.3.5).
Description of the method

In Section 4.3.5 we have considered the method of local descend: start at a coarse scale in the modified scale space and descend to finer scales only where a stopping criterion is not met. It is straightforward to apply this idea to the evaluation problem. We propose a method of multiscale evaluation that can be applied as follows. It is presumed that

- the deformation curve has been calculated,
- one of the well logs has been deformed accordingly,
- the new local covariance has been calculated,
- the modified scale space of one of the well logs has been determined.

The sequence of calculations starts at a coarse scale and we determine at this scale the mean value of the local covariance within each segment. In formula

\[ \tilde{C}_j^{(\sigma)} = \frac{1}{\tilde{x}_{j+1}^{(\sigma)} - \tilde{x}_j^{(\sigma)}} \int_{\tilde{x}_j^{(\sigma)}}^{\tilde{x}_{j+1}^{(\sigma)}} C(x) \, dx, \]

where we have used the notation introduced in Section 4.3. The stopping criterion is based on a threshold (\(\theta\)) on the mean value of the local covariance in a segment. If \(\tilde{C}_j^{(\sigma)} > \theta\) then local descend within segment \(j\) is terminated else local descend proceeds.

The stopping criterion can be extended with a minimum admissible layer thickness or minimum scale.

Remark 1
The final result will consist of segments of various lengths in which the average value of the local covariance exceeds the threshold and of small size segments where the average value is smaller than the threshold. In the results only the segments where the local covariance exceeds the threshold will be shown.

Remark 2
The value of the threshold can be selected freely. The histogram of the local covariance can give an indication what value to select.
Remark 3
We have required that a few outliers of the local covariance in a segment should not affect the results. This goal has partly been achieved by the use of a multiscale approach. Their influence can be further reduced if we do not determine the average value of local covariance within a segment but the median.

Two gamma-ray logs

In this section the multiscale evaluation method will be applied to the same gamma-ray logs as in Section 5.7.2. For the upper gamma-ray log in fig.5.7.1 we have determined the scale space and the modified scale space (see fig.5.7.10 and fig.5.7.11). The scale axis is discretized as

$$\sigma_k = 4 + (1.05)^k \quad k = 0, 1, \ldots$$

The threshold that is used is approximately the median of the local covariance (after deformation), which is the point for which the histogram has half of its area (right-hand side of fig.5.7.7) on either side. The threshold is

$$\theta = 0.73$$

In the experiment the minimum scale level has been set at of

$$\sigma_{\text{min}} = \sigma_{30} \approx 8.32$$

No other constraints are imposed.

In fig.5.7.12, the result of the multiscale evaluation of the two gamma-ray logs is shown. At the top of the picture the local covariance is shown (the same as in fig.5.7.2). The threshold is indicated by a straight line. In the middle and at the bottom of the picture the two not-deformed gamma-ray logs are depicted. Shading in the bar that separates the two logs indicate the segments where a good match has been achieved. It can be seen that it is possible to have segments of a good match, which have some bad outliers.

Two SP logs

In this section the multiscale evaluation method will be applied to the same SP logs as in Section 5.7.2. For the upper SP log in fig.5.7.3 we have determined the the modified scale space (see fig.5.7.13). The scale axis is discretized as

$$\sigma_k = 4 + (1.02)^k \quad k = 0, 1, \ldots$$
The median of the local covariance can be obtained from the histogram on the right-hand side of fig.5.7.8 and equals 0.61. The overall match of the two SP logs is worse than the match of the two gamma-ray logs. The value of the threshold is taken to be

$$\theta = 0.65.$$ 

A minimum scale has been set at

$$\sigma_{\text{min}} = \sigma_{73} \approx 8.24.$$ 

The result of the multiscale evaluation is depicted in fig.5.7.14. The outline of the picture is the same as fig.5.7.12. We can see that the multiscale evaluation method is very useful to detect the segments where a good match has been achieved.

**A SP log and an impedance log**

In this section the multiscale evaluation method will be applied to the same SP and impedance log as in Section 5.7.2. For the upper SP log in fig.5.7.5 we have determined the the modified scale space (see fig.5.7.15). The scale axis is discretized as

$$\sigma_k = 4 + (1.05)^k \quad k = 0, 1, \ldots$$

We already know that there is not much to be matched in the logs in this example. As a consequence the median of the local covariance has a low value. The threshold on the local covariance will be set at higher value than the median,

$$\theta = 0.70.$$ 

The minimum scale is

$$\sigma_{\text{min}} = \sigma_{30} \approx 8.32.$$ 

The result of the multiscale evaluation is depicted in fig.5.7.16. The outline of the picture is the same as fig.5.7.12. It can be seen that the method has been capable of detecting segments where similar log shapes can indeed be recognized.

**Remark on the threshold**

Although the histogram of the local covariance gives some indication the value of the threshold remains arbitrary. It is recommendable to compare the results for a number of thresholds instead of using only one.
5.7.4 Remarks

On the deformation

Due to faults it is possible that a bore hole intersects one and the same layer more than once. This effect known as repeated sequences constitutes a problem since in the preliminaries a causality condition has been imposed on the deformation curve: the deformation curve should be monotonically increasing. Repeated sequences lead to a violation of this condition, unless the number of times a certain layer is intersected is coincidentally the same for both well logs.

On interactive correlation

In practice well log correlation is often done visually or interactively. The well log analyst usually wants to mark a number points in the well log that, according to his experience and knowledge, belong to each other. A priori knowledge of the deformation curve of this kind can be incorporated straightforwardly in the matching method. It results in a considerable decrease in the number of calculations, because every known point of the deformation curve shuts out a part of rectangle in which the deformation curve has to be determined. Suppose the rectangle is given by the points

\[(x, y) \in [0, X] \times [0, Y].\]

A known point \((x_1, y_1)\) then shuts out the points for which

\[x > x_1, \; y < y_1 \text{ and } x < x_1, \; y > y_1.\]

5.8 Conclusions

A method for the matching of signals that are deformed with respect to each other has been derived and tested. The deformation may involve stretches, compressions and shifts and it has to satisfy a causality condition, that forbids inversions of parts of the signals. The matching is performed by minimization of the distance between the signals with respect to the deformation.

The distance measure determines if certain smooth additive and multiplicative variation influences the matching. A number of distance measures have been discussed. The local square error is insensitive to additive variation and the
local correlation measure is insensitive to multiplicative variation. The local covariance measure in insensitive to either of aforementioned variations, while the square error is sensitive to each of them.

By discretization of the integral expression in the distance measure, the minimization problem can be formulated as a dynamic programming problem. This means that a recursion formula can be derived, which reduces the complexity of the minimization to quadratic dependence on the number of data points.

Experiments have been done on synthetic data. In these synthetic examples the influence of the discretization width, steepness of the deformation curve and noise has been tested. The sensitivity of the method to additive and multiplicative variation in relation to the window length used in the local square error, local correlation measure and local covariance measure, has been considered. The results were good and in accordance with theoretical considerations. Theoretically, the method is able to cope with gaps and translations in the signals. However, in the experiments gaps in the signals appeared to constitute a major problem for the matching method. A solution to this will involve the subdivision of the signals as part of the matching procedure.

The matching method has also been tested on a pair of gamma-ray logs, on a pair of SP logs, and finally on a SP log with an impedance log. The results showed that edges, peaks and segments in one well log were matched to similar parts of the other well log, provided similarity was present. After deformation of one of the logs the local covariance appeared to have increased significantly.

In the last section a method has been introduced for the evaluation of the matching result. It is based on the multiscale segmentation of one of the well logs. The multiscale evaluation method provides the means to distinguish the segment where a good match has been achieved from segments where this is not the case. Outliers in the local covariance (or any other measure) within a segments do not significantly influence the result. This method is very useful for the recognition of similar log shapes.

It is concluded that signal matching by dynamic programming can be a very useful tool in well log correlation.
Figures Section 5.6

Figure 5.6.1: Recovered (solid) and model (dotted) deformation curve. The deformation curve relates the functions that are shown vertically and horizontally.

Figure 5.6.2: Difference between recovered and model deformation function: $m(s) - \hat{m}(s)$.

Figure 5.6.3: Difference between model deformation function and its representation on the grid: $m(s) - \hat{m}(s)$. 
Figure 5.6.4: Functions $g$ (solid) and $f$ (dashed), that are to be matched.

Figure 5.6.5: Functions $g$ (solid) and $\tilde{g}$ (dashed). $\tilde{g}$ has been obtained by deformation of $f$. 
Figure 5.6.6: Recovered (solid) and model (dotted) deformation curve. The deformation curve relates the functions that are shown vertically and horizontally.

Figure 5.6.7: Functions $g$ (solid) and $f$ (dashed), that are to be matched.

Figure 5.6.8: Functions $g$ and $\hat{g}$. $\hat{g}$ has been obtained by deformation of $f$. 
Figure 5.6.9: Smoothed recovered deformation curve (solid) and model (dotted) deformation curve. The deformation curve relates the functions that are shown vertically and horizontally.

Figure 5.6.10: Functions $g$ (solid) and $f$ (dashed), that are to be matched.

Figure 5.6.11: Functions $g$ and $\hat{g}$. $\hat{g}$ has been obtained by deformation of $f$. 
Figure 5.6.12: Error $E_m$ as a function of the number of samples $N$ for four different distance measures.

Figure 5.6.13: Error $E_m$ as a function of the number of samples $N$ for four different distance measures.

Figure 5.6.14: Influence of smoothing of the deformation function on the errors $E_m$ and $E_g$. 
Figure 5.6.15: Error $E_m$ as a function of half the window length ($d$), used in the calculation of the local square error, local correlation $m$, and local covariance $m$, for three values of $a$. 
Figure 5.6.16: Recovered (solid) and model (dotted) deformation curve. The deformation curve relates the functions that are shown vertically and horizontally.

Figure 5.6.17: Functions \( g \) (solid) and \( f \) (dashed), that are to be matched.

Figure 5.6.18: Functions \( g \) (solid) and \( \hat{g} \) (dashed). \( \hat{g} \) has been obtained by deformation of \( f \).
Figure 5.6.19: Error $E_v$ as a function of half the window length ($d$), used in the calculation of the local square error, local correlation $m.$ and local covariance $m.$, for two values of $N$. 
Figure 5.6.20: Influence of smoothing of the deformation function on the error $E_g$ for two values of $N$.

Figure 5.6.21: Recovered (solid) and model (dotted) deformation curve. The deformation curve relates the functions that are shown vertically and horizontally.
Figure 5.6.22: The noisy functions $\hat{g}$ (solid) and $\tilde{f}$ (dashed), that are to be matched.

Figure 5.6.23: Functions $g$ (solid) and $\hat{g}$ (dashed). $\hat{g}$ has been obtained by deformation of $f$.

Figure 5.6.24: Error $(E_m)^{\frac{1}{2}}$ as a function of the standard deviation $\sigma$ of the additive noise.

Figure 5.6.25: Error $(E_m)^{\frac{1}{2}}$ as a function of the standard deviation $\sigma$ of the additive noise.
Figure 5.6.26: Functions $g$ (solid) and $f$ (dashed), that are to be matched.

Figure 5.6.27: Recovered (solid) and model (dotted) deformation curve. The deformation curve relates the functions that are shown vertically and horizontally.
Figure 5.6.28: Functions $g$ (solid) and $\hat{g}$ (dashed). $\hat{g}$ has been obtained by deformation of $f$.

Figure 5.6.29: The difference $g(x) - \hat{g}(x)$. The result is a distorted sine.
Figure 5.6.30: Functions $g$ (solid) and $f$ (dashed), that are to be matched.

Figure 5.6.31: Recovered (solid) and model (dotted) deformation curve. The deformation curve relates the functions that are shown vertically and horizontally.
Figure 5.6.32: Functions \( g \) (solid) and \( \dot{g} \) (dashed). \( \dot{g} \) has been obtained by deformation of \( f \).

Figure 5.6.33: The difference \( g(x) - \dot{g}(x) \). The result is a distorted sine.

Figure 5.6.34: The error \( E_m \) as a function of \( d \) for several periods of the added sine. The amplitude of the sine is relatively large.

Figure 5.6.35: The error \( E_m \) as a function of \( d \) for several periods of the added sine. The amplitude of the sine is relatively small.
Figure 5.6.36: Functions $g$ (solid) and $f$ (dashed), that are to be matched.

Figure 5.6.37: Recovered (solid) and model (dotted) deformation curve. The deformation curve relates the functions that are shown vertically and horizontally.
Figure 5.6.38: Functions $g$ (solid) and $\hat{g}$ (dashed). $\hat{g}$ has been obtained by deformation of $f$.

Figure 5.6.39: The error $E_m$ as a function of $d$ for several periods of the sine by which $g$ has been multiplied.
Figure 5.6.40: Functions $g$ (solid) and $f$ (dashed), that are to be matched.

Figure 5.6.41: Recovered (solid) and model (dotted) deformation curve. The deformation curve relates the functions that are shown vertically and horizontally.
Figure 5.6.42: Functions $g$ (solid) and $\tilde{g}$ (dashed). $\tilde{g}$ has been obtained by deformation of $f$.

Figure 5.6.43: The error $E_n$ as a function of $d$ for several periods of the sines, that have been added to $g$ and by which $g$ has been multiplied.

Figure 5.6.44: Grey-value picture of the square-error matrix for a deformation curve with a shift.
Figure 5.6.45: Recovered (solid) and model (dotted) deformation curve. The deformation curve relates the functions that are shown vertically and horizontally.

Figure 5.6.46: Functions \( g \) (solid) and \( f \) (dashed), that are to be matched.

Figure 5.6.47: Functions \( g \) (solid) and \( \hat{g} \) (dashed). \( \hat{g} \) has been obtained by deformation of \( f \).
Figure 5.6.48: Recovered (solid) and model (dotted) deformation curve. The deformation curve relates the functions that are shown vertically and horizontally.

Figure 5.6.49: The square error between $f$ and $g$ along the recovered deformation curve (solid) and along the model deformation curve (dashed).
Figure 5.6.50: Recovered (solid) and model (dotted) deformation curve. The deformation curve relates the functions that are shown vertically and horizontally.

Figure 5.6.51: The square error between $f$ and $g$ along the recovered deformation curve (solid) and along the model deformation curve (dashed).
Figure 5.6.52: Recovered (solid) and model (dotted) deformation curve. The deformation curve relates the functions that are shown vertically and horizontally.
Figures Section 5.7

Figure 5.7.1: Two gamma-ray logs with local covariance (top).

Figure 5.7.2: Gamma-ray log (middle) and deformed gamma-ray log (bottom), with local covariance (top).
Figure 5.7.3: Two SP logs with local covariance (top).

Figure 5.7.4: SP log (middle) and deformed SP log (bottom), with local covariance (top).
Figure 5.7.5: SP log (middle) and Impedance log (bottom), with local covariance (top).

Figure 5.7.6: SP log (middle) and deformed Impedance log (bottom) with local covariance (top).
Figure 5.7.7: Histogram of the local covariance in fig.5.7.1 (left) and in fig.5.7.2 (right)

Figure 5.7.8: Histogram of the local covariance in fig.5.7.3 (left) and in fig.5.7.4 (right)

Figure 5.7.9: Histogram of the local covariance in fig.5.7.5 (left) and in fig.5.7.6 (right)
Figure 5.7.10: Gamma-ray log (top) and its scale space (bottom).

Figure 5.7.11: Gamma-ray log (top) and its modified scale space (bottom).
Figure 5.7.12: Two not-deformed gamma-ray logs with local covariance obtained after deformation (top). Segments with a good match are shaded.
Figure 5.7.13: SP log (top) and its modified scale space (bottom).

Figure 5.7.14: Two not-deformed SP logs with local covariance obtained after deformation (top). Segments with a good match are shaded.
Figure 5.7.15: SP log (top) and its modified scale space (bottom).

Figure 5.7.16: A SP log (middle) and a not-deformed impedance log with local covariance obtained after deformation (top). Segments with a good match are shaded.
Chapter 6

Concluding Remarks

The multiscale transform and the wavelet transform are very useful tools for signal and image processing. A considerable amount of theory on both transforms is available and many applications have been found, including data compression, time-variant filtering and segmentation. In this thesis we have added to both theory and application of these transforms.

A theory for the decomposition of discrete-time signals in discrete orthonormal wavelet bases has been developed. These wavelets are obtained by dilating, shifting and sampling a basic function called the wavelet function. A discrete multiscale transform is obtained by calculation of the inner product of the discrete-time signal with dilated, shifted and sampled versions of the so called scaling function.

It has been shown that there exists a strong relation with already existing techniques as subband (QMF-) filterbanks and Laplacian and Gaussian pyramids. In fact, under certain additional conditions, the wavelet function and the scaling function are uniquely determined by the QMF-filters of an exact reconstruction filterbank.

The multiscale transform has been used to develop a multiscale segmentation method for well logs. This segmentation method is better adjusted to the multiscale character of a well log than conventional techniques. The method is based on the behavior of edges across scales. Edge detection is done by determination of the zero crossings of the second derivative. In this way, edge detection at all scales in the multiscale transformed well log is equivalent with determination of the zero crossings of the wavelet transformed well log. If we impose the condition that a fine scale segment should not split up in smaller segments if scale becomes coarser, the Gaussian appears to be the only appropriate filter.

A signal matching method has been considered that is based on dynamic pro-
gramming. It determines the deformation between signals that have been shifted, stretched or compressed with respect to each other. By selection of a proper distance measure the matching can be made insensitive to smooth additive and/or multiplicative variations of the signals. The method is straightforward and requires no initial estimate of the solution. As for the application of the method, we have focussed on the matching of well logs. The results of tests on well logs as well as on synthetic data were very satisfactory. The multiscale segmentation method has successfully been applied for the determination of segments in the well logs where a good (or bad) match is present. This resulted in a multiscale evaluation method for the matching result.
References


Acknowledgement

The author wishes to thank Jos Alkemade at the Koninklijke/Shell Exploration and Production Laboratory in Rijswijk for four years of mentorship and pleasant cooperation.
Summary

A signal may exhibit a certain global behavior with a superposition of local features, for instance a trend with some high-frequent oscillations. It is also possible that the signal's frequency content varies considerably as a function of time. In both cases a transform that results in an expansion in basic functions with infinite duration, such as the Fourier transform, will give an inadequate description of the behavior of the signal. The multiscale transform and the wavelet transform have been devised to solve this particular problem in signal and image analysis.

Both the multiscale and the wavelet transform are so called affine transforms. In an affine transform a function is convoluted with a filter of which the width is controlled by a dilation parameter. The filter determines the specific properties of the transform.

(a) In the multiscale transform a low-pass (or smoothing) filter is used. The dilation parameter controls the amount of smoothing. The result is a description of the original signal at different levels of smoothness, which are called scales. The multiscale transform results in a separation between local and global behavior of a signal.

(b) In the wavelet transform a band-pass filter is used. The result is a description of the original signal in different frequency bands or, equivalently, a description of its time-varying spectrum. At each point and for each value of the dilation parameter the wavelet transform can be viewed as an inner product between the signal and a basic function, which is called a wavelet.

A considerable amount of theory has already been developed on the expansion of continuous-time signals in discrete sets of wavelets. That is, the construction of bases and frames of wavelets for the space of square-integrable functions $L^2(\mathbb{R})$. In this thesis we have extended the theory to the construction of
orthonormal wavelet basis for $l^2$, which means that discrete-time signals can be expanded in sets discrete-time wavelets. The discrete-time wavelets are obtained by shifting, dilating and subsequent sampling of the wavelet function. Part of the theory is the scaling transform which is in fact a discrete multiscale transform. The scaling transform is the inner product of the discrete-time signal with dilated, shifted and sampled versions of the scaling function.

It has been shown that the theory of discrete orthonormal wavelets is strongly connected to subband (QMF-) filterbanks. An exact reconstruction filterbank can be used for the efficient calculation of the wavelet expansion of a discrete-time signal. Moreover, the wavelet and scaling function uniquely determine the QMF-filters in the subband filterbank.

A multiscale segmentation method has been proposed that is based on the behavior of edges across scales. Edges correspond to extrema of the first derivative of a signal or, equivalently, zero crossings of its second derivative. The scale space of a signal has been defined as the set of edges of its multiscale transform. The scale space consequently consists of a set of zero crossings of the second derivative with respect to time of its multiscale transform, which is equivalent to the set of zero crossings of its wavelet transform.

In the time-scale plane the zero crossings form zero crossing curves which may vanish in pairs as scale increases. These curves give the position of an edge at a certain scale and point toward the position of the edge in the original signal. At any scale the original edge position can therefore be retrieved by coarse-to-fine tracking. After restoration of all original edge positions we obtain the modified scale space, which is a multiscale segmentation of the signal.

Since it is undesirable that fine scale segments split up as scale increases, no zero crossings may be created as scale increases. The Gaussian has appeared to be the only filter in the multiscale transform that can guarantee this property.

The multiscale segmentation method has been applied to the analysis of well logs. Well logs are measurements of a certain physical quantity obtained from a bore hole. Examples of physical quantities are the natural gamma radiation, the sound velocity and the specific density. We found that multiscale segmentation fits the multiscale character of the well logs, unlike conventional methods, and enables to perform well log analysis from a new point of view.

A signal matching method has been derived and tested for the matching of
signals that may be shifted, stretched and compressed with respect to each other. The matching involves the minimization of a distance measure between two signals with respect to the deformation. The deformation has to satisfy a causality condition that forbids inversion of parts of the signal. The distance measure determines whether certain multiplicative and additive variation on either of the signals influences the matching result. A number of these distance measures have been considered.

By discretization of an integral expression in the distance measure, the minimization problem can be formulated as a dynamic programming problem. This means that a recursion formula has been derived, which reduces the complexity of the minimization. Experiments on synthetic data have been performed, in which the influence of noise and multiplicative and additive variations has been examined.

The matching method has also been applied to the matching or correlation of well logs. Well logs that have been obtained from different bore holes in a certain area may have similar features but they can also exhibit differences due to different geological circumstances at different locations. For instance, dissimilar sedimentation rates may cause the well logs to be stretched or compressed with respect to each other. The results of the experiments showed that if similarity was present in the well logs the matching method was able to detect it, even if the two well logs represented different physical quantities.

As an extension to the matching method we have proposed the use of the multiscale segmentation method to evaluate the matching result. In this way, it is possible to determine the segments in the well logs where a good match has been achieved. These segments correspond to segments in the well logs themselves, they are of significant length and a few outliers do not affect the results. The multiscale evaluation method can be applied for the detection of similar log shapes and is a useful extension of the multiscale segmentation method.
Samenvatting

Een signaal kan een bepaald globaal gedrag vertonen met een superpositie van lokale verschijnselen, bijvoorbeeld een trend met hoogfrequente oscillaties. Ook kan het zijn dat de frequentie-inhoud van het signaal aanzienlijk varieert als functie van de tijd. In beide gevallen zal een transformatie die resulteert in een ontbinding in basisfuncties met een oneindige tijdsduur, zoals de Fourier-transformatie, niet een geschikte beschrijving van het gedrag van het signaal geven. De multiscale transformatie en de wavelet transformatie zijn bedoeld om dit probleem in de signaal- en beeldanalyse op te lossen.

Zowel de multiscale transformatie als de wavelet transformatie zijn zogenaamde affine transformaties. In een affine transformatie wordt een functie geconvolueerd met een filter waarvan de breedte wordt geregeld door een dilatatieparameter. Het filter bepaald de specifieke eigenschappen van de transformatie.

(a) In de multiscale transformatie wordt een laagdoorlatingsfilter (ook wel effeningsfilter) gebruikt. De dilatatieparameter regelt hoe sterk er gladgetrokken wordt. Het resultaat is een beschrijving van het oorspronkelijke signaal op verschillende gladheidsniveaus, die schalen worden genoemd. De multiscale transformatie resulteert in een scheiding tussen lokaal en globaal gedrag van een signaal.

(b) In de wavelet transformatie wordt een banddoorlatingsfilter gebruikt. Het resultaat is een beschrijving van het oorspronkelijke signaal in verschillende frequentiebanden of, equivalent daaraan, een beschrijving van zijn spectrum als functie van de tijd. In ieder punt en voor iedere waarde van de dilatatieparameter kan de wavelet transformatie worden gezien als het inwendig product tussen het signaal en een basisfunctie, die wavelet wordt genoemd.

Een aanmerkelijke hoeveelheid theorie is reeds ontwikkeld over de ontbinding van continue-tijd signalen in discrete verzamelingen van wavelets. Dat wil
zeggen, de constructie van bases en frames van wavelets voor de ruimte van kwadratisch integreerbare functies $L^2(\mathbb{R})$. In dit proefschrift is de theorie uitgebreid naar de constructie van orthonormale wavelet bases van $l^2$, hetgeen betekent dat discrete-tijd signalen kunnen worden ontbonden in verzamelingen van discrete-tijd wavelets.

De discrete-tijd wavelets worden verkregen door het verschuiven, dilateren en vervolgens bemonsteren van de wavelet functie. De schalingstransformatie vormt een onderdeel van de theorie en is in feite een discrete multiscale transformatie. Deze transformatie is het inwendig product van een discrete-tijd signaal met gedilateerde, verschoven en bemonsterde versies van de schalingsfunctie.

Er is aangetoond dat de theorie over discrete orthonormale wavelets sterk verbonden is met subband- (QMF-) filterbanken. Een exacte reconstructie filterbank kan worden gebruikt om de wavelet onttbinding van een discrete-tijd signaal efficiënt te berekenen. Bovendien worden de QMF-filters van de subband filterbank uniek bepaald door de wavelet functie en de schalingsfunctie.

Een multiscale segmentatie methode is voorgesteld die is gebaseerd op het gedrag van randen voor verschillende schalen. Randen corresponderen met extrema van de eerste afgeleide van het signaal ofwel met de nul-doorgangen van zijn tweede afgeleide. De scale space van een signaal is gedefinieerd als de verzameling randen van zijn multiscale getransformeerde. De scale space van een signaal bestaat derhalve uit een verzameling van nul-doorgangen van de tweede afgeleide naar de tijd van zijn multiscalegetransformeerde, hetgeen equivalent is aan de verzameling nul-doorgangen van zijn wavelet getransformeerde.

De nul-doorgangen vormen nul-doorgangscurves in het tijd-schaal vlak die paarsgewijs kunnen verdwijnen voor groter wordende schaal. Deze curves geven de positie van een rand op een een bepaalde schaal en wijzen naar de positie van de rand in het oorspronkelijke signaal. Derhalve kan op iedere schaal de oorspronkelijke positie van de rand worden teruggevonden door grof-naar-fijn te volgen. Na het herstellen van oorspronkelijke posities van de randen wordt de gemodificeerde scale space verkregen, hetgeen een multiscale segmentatie is van het signaal.

Aangezien het ongewenst is dat fijnschalige segmenten opsplitsen als de schaal groter wordt, mogen geen nieuwe nul-doorgangen ontstaan als de schaal groter wordt. De Gaussische functie bleek het enige filter te zijn in de multiscale transformatie dat deze eigenschap kan garanderen.
De multiscale segmentatie methode is toegepast voor de analyse van well logs. Well logs zijn metingen van een bepaalde fysische grootheid in een boorput. Voorbeelden zijn de natuurlijke gammastraling, de geluidssnelheid en de soortelijke dichtheid. Het blijkt dat de multiscale segmentatie methode zeer goed past bij het veelschappige karakter van de well logs, in tegenstelling tot conventionele methoden, en dat het de mogelijkheid biedt om well logs vanuit een nieuw gezichtspunt te analyseren.

Een signaal matching methode is afgeleid en getest voor het matchen van signalen die verschoven, uitgerekt of ingedrukt kunnen zijn ten opzichte van elkaar. Het matchen behelst het minimaliseren van een afstandsmaat tussen twee signalen met betrekking tot de vervorming. De vervorming moet aan een causaliteitsvoorwaarde voldoen die het verwisselen van stukken signaal verbiedt. De afstandsmaat bepaalt of zekere multiplicatieve en additieve variatie op de signalen het matchingresultaat beïnvloedt. Een aantal afstandsmaten is bekend.

Door discretisatie van een integraaluitdrukking in de afstandsmaat, kan het minimaliseringsprobleem worden geformuleerd als dynamisch programmeringsprobleem. Dit betekent dat een recursieformule is afgeleid, die de complexiteit van de minimalisering vermindert.

Experimenten, waarin de invloed van ruis en multiplicatieve en additieve variatie werd bestudeerd, zijn uitgevoerd op synthetische data.

De matching methode is ook toegepast voor het matchen of correleren van well logs. Well logs die zijn verkregen uit verschillende boorputten in een bepaald gebied hebben overeenkomstige kenmerken maar zij kunnen ook verschillen vertonen die te wijten zijn aan afwijkende geologische omstandigheden op de verschillende plaatsen. Ongelijke afzettingssnelheden kunnen er de oorzaak van zijn dat de well logs uitgerekt of ingedrukt zijn ten opzichte van elkaar. De resultaten van experimenten die zijn uitgevoerd waren goed. Als gelijkvormigheid aanwezig was in de well logs, was de matching methode in staat deze te detecteren, zelfs al representerden de well logs verschillende fysische grootheden.

Als uitbreiding van de matching methode is voorgesteld om de multiscale segmentatie methode te gebruiken voor de evaluatie van het matching resultaat. Op deze wijze is het mogelijk om segmenten te bepalen in de well logs waar een goede match is bereikt. Deze segmenten corresponderen met de segmenten
in de well logs zelf, zij hebben een significante lengte en een paar uitschieters hebben geen invloed op het resultaat. De multiscale evaluatie methode kan worden toegepast voor de detectie van overeenkomstige logvormen en is een nuttige uitbreiding van de multiscale segmentatie methode.
Curriculum Vitae