



Centrum voor Wiskunde en Informatica

REPORTRAPPORT

Automatic phase detection in seismic data using the discrete
wavelet transform

P.J. Oonincx

Probability, Networks and Algorithms (PNA)

PNA-R9811 October 1998

Report PNA-R9811
ISSN 1386-3711

CWI
P.O. Box 94079
1090 GB Amsterdam
The Netherlands

CWI is the National Research Institute for Mathematics and Computer Science. CWI is part of the Stichting Mathematisch Centrum (SMC), the Dutch foundation for promotion of mathematics and computer science and their applications.

SMC is sponsored by the Netherlands Organization for Scientific Research (NWO). CWI is a member of ERCIM, the European Research Consortium for Informatics and Mathematics.

Copyright © Stichting Mathematisch Centrum
P.O. Box 94079, 1090 GB Amsterdam (NL)
Kruislaan 413, 1098 SJ Amsterdam (NL)
Telephone +31 20 592 9333
Telefax +31 20 592 4199

Automatic Phase Detection in Seismic Data using the Discrete Wavelet Transform

P.J. Oonincx

CWI

P.O. Box 94079, 1090 GB Amsterdam, The Netherlands

ABSTRACT

Seismic data consist of traces, which contain information about a seismic event, but in some period of time the traces may be just noise. A trace which contains seismic information, is called a seismic signal. Seismic signals consist of several typically short energy bursts, called phases, exhibiting several patterns in terms of dominant frequency, amplitude and polarisation. Amongst others, a significant phase is the S-phase. We present a fast algorithm to detect the S-phase in a three-component seismic signal. This new approach is a combination of traditional S-phase detection methods from seismology and the discrete wavelet transform. Stability and correctness of the algorithm will be proved and results will be presented to demonstrate the algorithm.

1991 Mathematics Subject Classification: 42C15, 65G99, 86A15, 94A12, 94A13.

Keywords and Phrases: Wavelets, seismology, feature detection, phase pickers, cross-power matrix, energy distribution.

Note: Work carried out under project PNA 4.2 "Wavelets" and supported financially by the Technology Foundation (STW), project no. CWI44.3403.

1. INTRODUCTION

A seismic signal consists of several different phases (waves), which characterise the type of the considered seismic signal. Amongst others, significant phases are the P-phase and the S-phase, which we consider here. The problem we are dealing with is to detect the arrival time of the S-phase if the P-phase arrival time is already known. An accurate determination of these arrival times is important for detecting the location and the type of a seismic event.

The S-phase arrival time is determined in a three-component seismogram, representing motion on a ground detector in three mutually orthogonal directions, two in the horizontal plane (x-y plane) and one vertical direction (z-axis). An example of a three-component seismogram is depicted in Figure 1. The detection of the S-phase arrival time is based on some physical differences between the P-phase and the S-phase, as described thoroughly in [1]. The first property we use is the delay in arrival time. When a seismic event takes place the S-phase arrives after the P-phase at the earth's surface. Furthermore the S particle motion, i.e. the direction of the S-phase when it arrives at the earth's surface, is contained in a plane perpendicular to the direction of the P particle motion, called the S-plane. The P-phase travels along the travel direction of the seismic event. Finally, comparing the frequency spectra of both phases, the P-phase appears at higher frequencies than the S-phase.

The automatic S-phase detection algorithm that we present in this paper is a combination of traditional methods to detect S-phases as described before and the discrete wavelet transform. This transform analyses the three components of the seismogram at several scales (frequency bands). The idea to

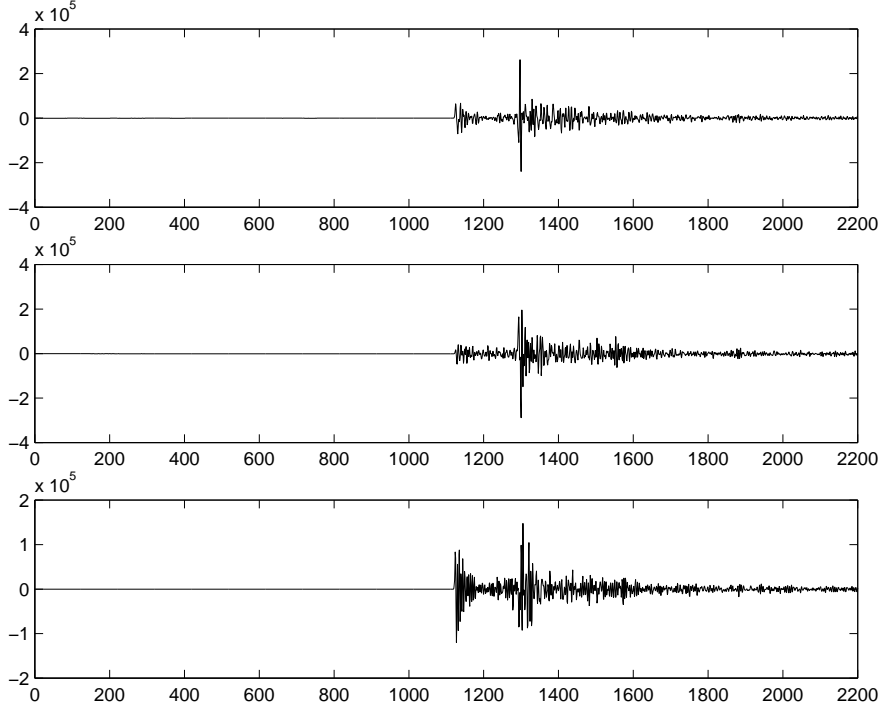


Figure 1: A three-component seismogram

analyse the components in this way has been described already in the literature, e.g. [2, 15]. However in these papers only the wavelet transform itself has been used as a phase detector in seismograms. The traditional methods we use in our approach are very well known in seismology and have been discussed in the past in various papers, e.g. [5, 10, 18]. Results of this new algorithm are compared with results based on ideas of Cichowicz [5].

2. CHARACTERISTIC FUNCTIONS FOR PHASE PICKING

In this section we discuss existing methods for phase picking, which are nowadays commonly used. These methods are based on statistical and physical properties of the seismogram, but do not deal with its scaling behaviour, which is made explicit by the wavelet transform.

A common strategy to detect phase arrival times is to construct one or more so-called characteristic functions. These are discrete-time functions, with some specific properties at the time samples, where the phases arrive. Here we discuss some of them both for detecting P-phase arrival times and S-phase arrival times.

2.1 STA/LTA

We consider a sampled one-component seismic signal $u \in l^1(\mathbb{N})$, with $\mathbb{N} = \{0, 1, 2, \dots\}$. With this sequence we can build a characteristic function given by

$$\kappa(i) = u(i)^2 + C_u(i) \cdot (u(i) - u(i-1))^2, \quad (2.1)$$

for $i \geq 1$, with $C_u(i)$ some time-dependent parameter, depending continuously on $u(0), \dots, u(i)$. Further we define $\kappa(0) = u(0)^2$. With (2.1) the short time average (STA) and the long time average (LTA) of the signal u are defined recursively by

$$STA(i) = C_1 \kappa(i) + (1 - C_1) STA(i - 1), \quad (2.2)$$

$$LTA(i) = C_2 \kappa(i) + (1 - C_2) LTA(i - 1), \quad (2.3)$$

for $i \geq 1$, and with initial values $STA(0) = LTA(0) = \kappa(0)$. The values of the constants $0 < C_1 < 1$ and $0 < C_2 \ll 1$ depend on the kind of seismic event. We observe, that for large numbers $C_1 \sim 1$, the STA measures the local behaviour of the signal. For small numbers $C_2 \sim 0$, the LTA measures the global behaviour of the signal. To get some insight in the values of C_1 and C_2 in practice, Baer and Kradolfer used in their paper [3] for C_1 the constants 0.67 for local events and 0.65 for teleseismic events. In the same paper, for C_2 the values 0.015 and 0.008 are used for local and teleseismic events respectively. The signals are sampled at a sampling rate of 64 Hz in [3].

The arrival time i_P of the first event, i.e. the P-phase arrival time, is declared as

$$i_P = \inf\{i \in \mathbb{N} \mid STA(i)/LTA(i) > C_3\}, \quad (2.4)$$

for some constant $C_3 > 1$. In [3] choices for local and teleseismic events are $C_3 = 9$ and $C_3 = 10$ respectively.

This method yields a first approximation of the P-phase arrival time. Clearly it is based on sudden increases of high frequencies in the signal. Therefore, results by Baer and Kradolfer were satisfactory for local events. However, the Allen picker failed to detect many P arrival times of teleseismic signals, which show an increase in amplitude rather than in frequency at arrival times, see [3]. Another shortcoming is the use of a hard threshold in (2.4). An adaptive threshold would be more convenient to pick arrival times in signals with a low signal-to-noise rate (SNR). To avoid these problems, Baer and Kradolfer developed a more sophisticated algorithm. Since the idea is similar to the idea behind Allen's picker, here we only refer to their paper [3] on this subject.

2.2 Energy ratio

In dealing with three-component seismic data, a characteristic function can be built to detect coherent seismic motion in a 2-dimensional plane. We consider a three-component seismic signal $\underline{u} = (u_1, u_2, u_3) \in l^2(\mathbb{N}, \mathbb{R}^3)$. The ratio of the energy in the horizontal plane and the total energy at some time sample i is given by

$$\kappa(i) = \frac{\sum_{n=i}^{i+N-1} (u_1(n)^2 + u_2(n)^2)}{\sum_{n=i}^{i+N-1} (u_1(n)^2 + u_2(n)^2 + u_3(n)^2)}, \quad (2.5)$$

for some $N \in \mathbb{N}^+$. Here N is some window length, depending on the sample rate and the kind of seismic signal. So κ tends to 1 if most of the seismic energy is concentrated in the xy -plane, while κ is small if most of the seismic energy is concentrated in the vertical direction.

This characteristic function can be used in a more sophisticated way, if the direction of the P particle motion, i.e. the direction of the P-phase when it arrives at the earth's surface, is known. Since it is

known that the S-phase arrives in a direction perpendicular to the direction of the P particle motion, we can use the ratio of energy in the transversal plane (the plane perpendicular to the direction of the P particle motion) to the total energy in the signal. How this can be done is explained in Section 3.

2.3 Characteristic Functions based on a Cross-Power Matrix

Again we consider a three-component seismic signal $\underline{u} \in l^2(\mathbb{N}, \mathbb{R}^3)$. Now we present some characteristic functions for three-component seismic data for both detecting P arrival times and detecting S arrival times. All of these characteristic functions are based on the time-dependent cross-power matrix, defined as follows.

Definition 2.1

Let $N \in \mathbb{N} \setminus \{0\}$ and $u_1, u_2, u_3 \in l^2(\mathbb{N})$. Then the N -point cross-power matrix of u_1, u_2 and u_3 at $i \in \mathbb{N}$ is given by

$$M(i) = \begin{pmatrix} \langle u_1, u_1 \rangle_i & \langle u_1, u_2 \rangle_i & \langle u_1, u_3 \rangle_i \\ \langle u_2, u_1 \rangle_i & \langle u_2, u_2 \rangle_i & \langle u_2, u_3 \rangle_i \\ \langle u_3, u_1 \rangle_i & \langle u_3, u_2 \rangle_i & \langle u_3, u_3 \rangle_i \end{pmatrix}, \quad (2.6)$$

for $i \in \mathbb{N}$. Here $\langle u_n, u_m \rangle_i, i \in \mathbb{N}$, denotes the semi-inner product given by $\langle u_n, u_m \rangle_i = 1/N \sum_{k=i}^{i+N-1} u_n(k) \cdot u_m(k)$. As discussed already in Section 2.2, the window length N depends on the sample rate and the kind of seismic event.

In practice, we first create a signal \underline{u} , with zero mean at each component, by subtracting the means of the components from the seismic signal. The reason for doing this is to neglect possible offsets without seismic cause. These can be generated by the measurement equipment.

We observe that $M(i)$ is the Gram matrix of the set $\{\chi_{[i, i+N-1]} u_1, \chi_{[i, i+N-1]} u_2, \chi_{[i, i+N-1]} u_3\}$, with χ_X denoting the characteristic function on \mathbb{N} of a discrete set X . So $M(i)$ is a positive semi-definite matrix and therefore the eigenvalues of $M(i)$ are real and positive or zero. Furthermore the eigenvectors can be chosen to form an orthonormal basis in \mathbb{R}^3 .

We briefly discuss some characteristic functions based on the cross-power matrix defined in (2.6). When discussing these functions, we assume that $M(i) \neq 0$, i.e. at least one of the three components is not equal to 0, and that all eigenvalues of $M(i)$ have multiplicity 1, which are quite realistic assumptions. More information on these functions can be found in [6, 10, 11, 13, 18].

2.3.1 Rectilinearity: Let $\lambda_1(i) \geq \lambda_2(i) \geq \lambda_3(i) \geq 0$ be the eigenvalues of the cross-power matrix $M(i)$ of the seismic signal \underline{u} . Following [13] the rectilinearity function κ is defined by

$$\kappa(i) = \frac{\lambda_1(i) - \lambda_2(i)}{\lambda_1(i)}. \quad (2.7)$$

This function measures the degree of linear polarisation of the three-component signal, i.e. κ increases if more coherent seismic energy is located in one principal direction. Since the P-phase is (nearly) linearly polarised, we may expect that this function will attain its maximal value at i_P , the P-phase arrival time. Therefore the rectilinearity function can be used to detect P-phase arrival times. One may define other rectilinearity functions in terms of the eigenvalues of $M(i)$, see e.g. [11]. However,

we restrict ourselves to the function in (2.7).

2.3.2 Planarity: Let $\lambda_1(i) \geq \lambda_2(i) \geq \lambda_3(i) \geq 0$ be as defined before. Following [11], the planarity function κ is defined by

$$\kappa(i) = \frac{\lambda_1(i) + \lambda_2(i) - 2\lambda_3(i)}{\lambda_1(i) + \lambda_2(i)}. \quad (2.8)$$

This function measures the degree of polarisation of the three-component signal in a plane. Since the S-phase particle motion should be confined to a plane, we may expect that this function will attain its maximal value at i_S , the S-phase arrival time. Therefore the rectilinearity function is an appropriate tool to detect S-phase arrival times.

2.3.3 Deflection angle: Let $\underline{v}_1(i)$ denote the eigenvector of $M(i)$ corresponding to $\lambda_1(i)$, with $\lambda_1(i)$ as defined before. This eigenvector $\underline{v}_1(i)$ represents the direction of the particle motion at time i . We define the direction of the P particle motion as $\underline{v}_1(i_P)$. Then the deflection angle is defined by

$$\kappa(i) = \frac{2}{\pi} \arccos \left(\frac{|(\underline{v}_1(i), \underline{v}_1(i_P))|}{\|\underline{v}_1(i)\| \cdot \|\underline{v}_1(i_P)\|} \right). \quad (2.9)$$

Since we assumed $|\sigma(M(i))| = 3$, i.e. the spectrum consists of three different eigenvalues κ is uniquely determined for all $i \in \mathcal{N}$. Note that $\kappa(i_P) = 0$. Furthermore, since the direction of the S-phase particle motion is perpendicular to $\underline{v}_1(i_P)$, κ attains its maximum 1 at the S arrival time i_S .

2.3.4 Degree of polarisation: Let $\lambda_1(i) \geq \lambda_2(i) \geq \lambda_3(i) \geq 0$ be as defined before. Following [18], the degree of polarisation is defined by

$$\kappa(i) = \frac{(\lambda_1(i) - \lambda_2(i))^2 + (\lambda_1(i) - \lambda_3(i))^2 + (\lambda_2(i) - \lambda_3(i))^2}{2 \cdot (\lambda_1(i) + \lambda_2(i) + \lambda_3(i))^2}. \quad (2.10)$$

This characteristic function can be used both for detecting P-phase arrival times and S-phase arrival times, since all types of seismic polarisation exhibit a high degree of polarisation. This means that we may expect maxima for κ at both i_P and i_S .

Note that all presented κ depend on the cross-power matrix M in (2.6), and therefore they depend on the window length N . Later in this paper we will also discuss window lengths that depend on the kind of analysed signal. Also we observe that for all characteristic functions κ introduced in this section we have

$$\forall_{i \in \mathcal{N}} \quad 0 \leq \kappa(i) \leq 1. \quad (2.11)$$

Further, all κ as described in this section will ideally attain their maxima at at least one of the phase arrival times. However to detect S-phase arrival times we are only interested in characteristic functions that can reach their maximal value at i_S . Therefore (2.7) is not useful in our approach. Also, from the definitions of the characteristic functions in this section it is not obvious whether a maximum of κ is

related to i_S or to another feature of the seismic signal. Therefore we use a combination of $M > 0$ characteristic functions $\kappa_1, \dots, \kappa_M$, such that

$$\prod_{m=1}^M \kappa_m^2(i) = \max_{n \in \mathbb{N}} \prod_{m=1}^M \kappa_m^2(n) \iff i = i_S. \quad (2.12)$$

We use the square product of several characteristic functions to emphasize the maximal values attained in each function at i_S and to reduce other (local) maxima, related to features in the signal other than the S-phase arrival. So our aim is to choose characteristic functions κ_m , such that only the S-phase arrival will give raise to a maximum in all chosen κ_m . In this section we have already met some candidates for this product of functions. In the next section we will slightly change (2.5), such that it also becomes an appropriate candidate.

3. SOME REMARKS ON PHASE PICKERS IN PRACTICE

In this section we discuss how the phase picking methods can be used in practice in an efficient way. Moreover, we discuss a method to transform the signal into longitudinal and transversal directions and we study properties of the cross-power matrix as introduced in (2.6). In particular we discuss the eigenvalue and eigenvector behaviour of $M(i)$. We will see that these behaviours yield stable characteristic functions κ , i.e. small changes in the cross-power matrix result in small changes in the characteristic functions. Further, we discuss problems that appear when we analyse data coming from seismic events. With some examples it will be clear why the discrete wavelet transform can be used to avoid or to reduce these problems.

3.1 Longitudinal and transversal directions

We discuss how we can transform three-component seismic data measured along the standard basis vectors in \mathbb{R}^3 into a three-component signal with components in the longitudinal direction and two transversal directions. The longitudinal direction is the direction of the P particle motion. The transversal directions are mutually orthogonal and are chosen in the plane perpendicular to the longitudinal direction. This transversal plan is also called the S-plane, since the direction of the S particle motion is in the S-plane.

We consider the three-component seismic signal $\underline{u} = (u_1, u_2, u_3)^T \in l^2(\mathbb{N}, \mathbb{R}^3)$. Now assume we have detected the P arrival time with some technique as discussed in the previous section. In the case we use a detection method for one-component signals, we only use the vertical component u_3 of \underline{u} to detect the P arrival time. Using one component instead of all three components will not cause any time delay, only the amplitude of the signal is affected. This is caused by projecting the signal orthonormally onto the three axes. The choice for u_3 lies in the fact that the seismic waves coming from their source often reach the earth's surface in a direction close to the vertical direction.

We denote by i_P the detected P arrival time. At this arrival time we compute the cross-power matrix as defined in (2.6). From this matrix, we compute the eigenvalues $\lambda_1(i_P) \geq \lambda_2(i_P) \geq \lambda_3(i_P) \geq 0$ and their corresponding mutually orthonormal eigenvectors $\underline{v}_1(i_P), \underline{v}_2(i_P)$ and $\underline{v}_3(i_P)$. These eigenvalues indicate the distribution of seismic energy along three directions, given by their corresponding eigenvectors. Now, the direction of polarisation at i_P is the longitudinal direction, the direction corresponding to the largest amount of seismic energy $\lambda_1(i_P)$, given by \underline{v}_1 . Orthogonal to the longitudinal direction, we find two transverse directions \underline{v}_2 and \underline{v}_3 . These two eigenvectors span together the S-plane.

We can use the new directions to transform the ground motion $\underline{u}(i) = u_1(i)\underline{e}_1 + u_2(i)\underline{e}_2 + u_3(i)\underline{e}_3$ into motion along the longitudinal and the transversal directions ($\underline{v}_1(i_P), \underline{v}_2(i_P)$ and $\underline{v}_3(i_P)$). Here $\underline{e}_1, \underline{e}_2, \underline{e}_3$ denote the standard basis vectors in \mathbb{R}^3 . This transformation is given by applying the orthonormal matrix $V(i_P)$ on $\underline{u}(i)$, where $V(i_P)$ is given by

$$V(i_P) = \left(\underline{v}_1(i_P) \mid \underline{v}_2(i_P) \mid \underline{v}_3(i_P) \right)^T. \quad (3.1)$$

The characteristic functions of Section 2 can now be applied to the transformed signal. This can lead to characteristic functions, which are more focused on the S arrival time detection, as we can see in the following example.

We consider the characteristic function (2.5) representing the fraction of energy in two directions to the total amount of energy in the three-component signal. Using the transformed signal for this κ we get

$$\kappa(i) = \frac{\sum_{n=i}^{i+N-1} (u_Q(n)^2 + u_T(n)^2)}{\sum_{n=i}^{i+N-1} (u_L(n)^2 + u_Q(n)^2 + u_T(n)^2)}, \quad (3.2)$$

for some $N \in \mathbb{N} \setminus \{0\}$. Here u_L, u_Q and u_T denote the longitudinal and two transversal directions respectively, which are given by

$$\begin{pmatrix} u_L(i) \\ u_Q(i) \\ u_T(i) \end{pmatrix} = V(i_P) \begin{pmatrix} u_1(i) \\ u_2(i) \\ u_3(i) \end{pmatrix}. \quad (3.3)$$

Now κ denotes the fraction of energy in the S-plane to the total amount of energy in the signal. So κ reaches a maximal value if all energy is located in the S-plane, which is ideally the case at the S-phase arrival time. Further this energy ratio satisfies (2.11). These two properties make also this characteristic function (3.2) an appropriate candidate for the product function as described in (2.12).

At the end of this subsection, we consider what happens if we compute the characteristic functions of Section 2 for a transformed 3-component signal. Particularly, we consider the error in these functions, due to an incorrect determination of the P arrival time.

Lemma 3.1

Let $u_1, u_2, u_3 \in l^2(\mathbb{N})$ and $M(i)$ the cross-power matrix of u_1, u_2 and u_3 at $i \in \mathbb{N}$ given by (2.6). Further, let V be a (3×3) matrix. Then $VM(i)V^T$ is the cross-power matrix of $(V\underline{u})_1, (V\underline{u})_2$ and $(V\underline{u})_3$ at $i \in \mathbb{N}$, where $\underline{u} = (u_1, u_2, u_3)^T$.

Proof

We write $M(i) = 1/N \sum_{k=i}^{i+N-1} \underline{u}(k)\underline{u}(k)^T$, According to this notation, we can write the cross-power

matrix of $(V\underline{u})_1, (V\underline{u})_2$ and $(V\underline{u})_3$ at $i \in \mathbb{N}$ as

$$\begin{aligned} M'(i) &= 1/N \sum_{k=i}^{i+N-1} V\underline{u}(k) (V\underline{u}(k))^T, \\ &= 1/N \sum_{k=i}^{i+N-1} V\underline{u}(k)\underline{u}(k)^T V^T = VM(i)V^T. \end{aligned}$$

□

From this lemma it follows, that $\sigma(M(i)) = \sigma(M'(i))$ if V is orthonormal and that the eigenvectors of $M'(i)$ are given by Vu_k , $k = 1, 2, 3$, with u_k the eigenvectors of $M(i)$. Obviously the characteristic functions based on the eigenvalues of the cross-power matrix do not change after transformation (3.3), since $\sigma(M(i)) = \sigma(M'(i))$. Since this holds for any orthonormal transformation, these functions also keep unchanged after transforming with the orthonormal matrix $V(i'_P)$, where i'_P is the incorrect determined P arrival time. To resume, the characteristic functions (2.7), (2.8) and (2.10) remain the same for all $i \in \mathbb{N}$ after a orthonormal transformation of the three-component signal.

In order to analyse what happens to the characteristic functions (2.9) and (3.2) if the P arrival time is determined incorrectly, we derive some properties of these functions. For this, we use i'_P to denote an incorrect determined P arrival time and u'_L, u'_Q and u'_T as the signal components after transforming the signal with $V(i'_P)$.

Lemma 3.2

Let κ be as defined in (3.2) and let κ' be as defined in (3.2) with u'_L, u'_Q and u'_T substituted for u_L, u_Q and u_T respectively. Then

$$\|\kappa - \kappa'\|_\infty \leq \|\underline{v}_1(i_P)\underline{v}_1(i_P)^T - \underline{v}_1(i'_P)\underline{v}_1(i'_P)^T\|_2, \quad (3.4)$$

with $\|\cdot\|_2$ the operator norm on \mathbb{R}^3 .

Proof

First we write (3.2) as

$$\begin{aligned} \kappa(i) &= \sum_{n=i}^{i+N-1} (QV(i_P)\underline{u}(n), V(i_P)\underline{u}(n)) / \sum_{n=i}^{i+N-1} (V(i_P)\underline{u}(n), V(i_P)\underline{u}(n)) \\ &= \sum_{n=i}^{i+N-1} (V(i_P)^T QV(i_P)\underline{u}(n), \underline{u}(n)) / \sum_{n=i}^{i+N-1} (\underline{u}(n), \underline{u}(n)), \end{aligned}$$

with Q the orthonormal projection onto $\{\underline{e}_2, \underline{e}_3\}$, and in the same manner we get

$$\kappa'(i) = \sum_{n=i}^{i+N-1} (V(i'_P)^T QV(i'_P)\underline{u}(n), \underline{u}(n)) / \sum_{n=i}^{i+N-1} (\underline{u}(n), \underline{u}(n)).$$

So

$$\begin{aligned}
& |\kappa(i) - \kappa'(i)| \\
&= \left| \frac{\sum_{n=i}^{i+N-1} ((V(i_P)^T QV(i_P) - V(i'_P)^T QV(i'_P))\underline{u}(n), \underline{u}(n))}{\sum_{n=i}^{i+N-1} (\underline{u}(n), \underline{u}(n))} \right| \\
&\leq \frac{\sum_{n=i}^{i+N-1} |((V(i_P)^T QV(i_P) - V(i'_P)^T QV(i'_P))\underline{u}(n), \underline{u}(n))|}{\sum_{n=i}^{i+N-1} (\underline{u}(n), \underline{u}(n))} \\
&\leq \|V(i_P)^T QV(i_P) - V(i'_P)^T QV(i'_P)\|_2 \frac{\sum_{n=i}^{i+N-1} (\underline{u}(n), \underline{u}(n))}{\sum_{n=i}^{i+N-1} (\underline{u}(n), \underline{u}(n))} \\
&= \|V(i_P)^T PV(i_P) - V(i'_P)^T PV(i'_P)\|_2,
\end{aligned}$$

with $P = I - Q$. A straightforward calculation yields $V(i_P)^T PV(i_P) = \underline{v}_1(i_P)\underline{v}_1(i_P)^T$ and $V(i'_P)^T PV(i'_P) = \underline{v}_1(i'_P)\underline{v}_1(i'_P)^T$. Further, since this upper bound holds for any $i \in \mathcal{I}N$, we can also take the supremum of $|\kappa(i) - \kappa'(i)|$ over all $i \in \mathcal{I}N$, which completes the proof. \square

We observe that the upper bound in (3.4) is sharp, so that it is a good indication for the error in the ratio of transversal to total energy, due to an incorrect P-phase arrival time determination.

The deflection angle (2.9) is also affected by using an incorrect determination of the P-phase arrival time. To get some insight how (2.9) is affected by this error in the P-phase arrival time, we rewrite (2.9) by using $\underline{v}_1(i_P) = V(i_P)^T \underline{e}_1$. Further we assume that the eigenvectors of $M(i)$ are chosen to be orthonormal. So

$$\begin{aligned}
\kappa(i) &= 2 \arccos (|(\underline{v}_1(i), \underline{v}_1(i_P))|) / \pi = 2 \arccos (|(\underline{v}_1(i), V(i_P)^T \underline{e}_1)|) / \pi \\
&= 2 \arccos (|(V(i_P)\underline{v}_1(i))_1|) / \pi.
\end{aligned}$$

We observe that $(V(i_P)\underline{v}_1(i))_1$ is the first component of the first eigenvector of $V(i_P)M(i)V(i_P)^T$, the cross-power matrix of the transformed three-component signal. In the same manner we get $\kappa'(i) = 2 \arccos (|(V(i'_P)\underline{v}_1(i))_1|) / \pi$. So the error in the deflection angle due to an incorrect P arrival time determination is given by

$$|\kappa(i) - \kappa'(i)| = 2 \left| \arccos (|(V(i_P)\underline{v}_1(i))_1|) - \arccos (|(V(i'_P)\underline{v}_1(i))_1|) \right| / \pi, \quad (3.5)$$

for all $i \in \mathcal{I}N$.

Note that we did not discuss (2.1) in this section, since the STA/LTA-method is a general signal processing tool with no special relation to the physical structure of the three-component seismic signal. Of course this method can also be used on the transverse directions of the transformed three-component signal to detect the S arrival time. However it is not obvious that the STA/LTA ratio will not reach a certain threshold if the signal is not transformed correctly.

3.2 Stability of the characteristic functions

In the previous subsection we considered the influence of determining an incorrect P-phase arrival time to the characteristic functions of Section 2. Here we will discuss the stability of these functions

in the sense that small changes in the signals u and \underline{u} or small errors in the computations of the cross-power matrices will lead to small changes in the characteristic functions. By definition it is clear, that the characteristic functions defined in (2.1) and (2.5) are stable. Therefore we only consider the stability of the characteristic functions of Section 2.3.

Before we derive results on the stability of these functions, we present some auxiliary results from perturbation theory.

Theorem 3.3

Let $M = A + \Delta$, with M, A and Δ Hermitian ($n \times n$) matrices. Let further $\mu_1 \geq \dots \geq \mu_n$, $\alpha_1 \geq \dots \geq \alpha_n$ and $\delta_1 \geq \dots \geq \delta_n$ be the eigenvalues of M, A and Δ respectively. Then

$$\alpha_k + \delta_n \leq \mu_k \leq \alpha_k + \delta_1. \quad (3.6)$$

Proof

Cf. [9]

□

Corollary 3.4

Let \mathcal{H}_n denote the set of all Hermitian ($n \times n$) matrices and let $\lambda : \mathcal{H}_n \rightarrow \mathbb{R}^n$ be given by $\lambda(H) = (\lambda_1, \dots, \lambda_n)$, with $\lambda_1 \geq \dots \geq \lambda_n \geq 0$ the eigenvalues of $H \in \mathcal{H}_n$. Then λ is continuous on \mathcal{H}_n .

Proof

Let $\varepsilon > 0$ and let Δ be a Hermitian ($n \times n$) matrix, with eigenvalues $\delta_1 \geq \dots \geq \delta_n$ and such that $\|\Delta\|_2 < \varepsilon$. Further, let $\lambda'_1 \geq \dots \geq \lambda'_n$ denote the eigenvalues of $H + \Delta$. Then by (3.6)

$$|\lambda'_k - \lambda_k| \leq \max\{|\delta_1|, |\delta_n|\} \leq \|\Delta\|_2 < \varepsilon.$$

Therefore $\|\lambda(H + \Delta) - \lambda(H)\|_2 < \varepsilon$.

□

From Corollary 3.4 the following result follows immediately.

Corollary 3.5

Let $f : \mathbb{R}^n \rightarrow [0, 1]$ be continuous and let \mathcal{H}_n denote the set of all Hermitian ($n \times n$) matrices. Then $\kappa = f \circ \lambda$ is continuous on \mathcal{H}_n , with λ as defined in Corollary 3.2

Since the characteristic functions in (2.7), (2.8) and (2.10) can be written as described in the previous corollary, they are stable. So small changes in the cross-power matrix (2.6) will yield small changes in these characteristic functions.

To prove the stability of (2.9), we use another result from perturbation theory, which is proved in [9].

Theorem 3.6

Let $M = A + \Delta$, with M, A and Δ ($n \times n$) matrices. Further, let $\alpha_1, \dots, \alpha_n$ be the eigenvalues of A and let $\underline{u}_1, \dots, \underline{u}_n$ be their corresponding eigenvectors. Finally, we assume that all eigenvalues have

multiplicity 1. Then the eigenvectors $\underline{v}_1, \dots, \underline{v}_n$ of M can be written as

$$\underline{v}_i = \underline{u}_i + \sum_{l \neq i} \frac{(\Delta \underline{u}_i, \underline{w}_l)}{(\alpha_i - \alpha_l)(\underline{u}_l, \underline{w}_l)} \underline{u}_l, \quad (3.7)$$

with $\underline{w}_1, \dots, \underline{w}_n$ the eigenvectors corresponding to the eigenvalues $\overline{\alpha}_1, \dots, \overline{\alpha}_n$ of A^* .

With this theorem we can prove the following corollary.

Corollary 3.7

Let H be a Hermitian ($n \times n$) matrix, $n > 1$, with eigenvectors $\underline{u}_1, \dots, \underline{u}_n$, $\|\underline{u}_k\| = 1$, $\underline{u}_k(1) \geq 0$, $k = 1, \dots, n$. Further we assume that all eigenvalues have multiplicity 1. Then the function $g_k(H) = \underline{u}_k$, $k = 1, \dots, n$ is continuous.

Proof

Let $\varepsilon > 0$ and let Δ be a ($n \times n$) matrix, $n > 1$, with $\|\Delta\|_2 < \varepsilon \cdot \alpha / (n - 1)$ with $\alpha = \min_{l \neq k} |\alpha_k - \alpha_l|$. Here $\alpha_1, \dots, \alpha_n$ are the eigenvalues of H . Then by (3.7)

$$\begin{aligned} \|g_k(H + \Delta) - g_k(H)\|_2 &= \left\| \sum_{l \neq k} \frac{(\Delta \underline{u}_k, \underline{u}_l)}{\alpha_k - \alpha_l} \underline{u}_l \right\|_2 \\ &\leq \sum_{l \neq k} \left| \frac{(\Delta \underline{u}_k, \underline{u}_l)}{\alpha_k - \alpha_l} \right| \\ &\leq (n - 1) \|\Delta\|_2 / \alpha < \varepsilon. \end{aligned}$$

□

Particularly Corollary 3.7 yields, that $g_1(M(i))$ is continuous, with $M(i)$ the cross-power matrix as defined in (2.6), under the assumption that $\lambda_1(i) \neq \lambda_2(i)$. Since this means, that the characteristic function in (2.9) is a composition of continuous functions, obviously also characteristic function (2.9) is stable if $|\sigma(M(i))| = 3$.

At this moment we have proved that the characteristic functions (2.7), (2.8), (2.9) and (2.10) depend continuously on the matrices $M(i)$. Obviously, changes in these matrices are not only due to computational errors, but generally the matrices $M(i)$ will vary due to the variation in the seismic signal \underline{u} . Therefore we will show that the cross-power matrices depend continuously on \underline{u} in the following sense.

Lemma 3.8

Let $\underline{u} \in l^2(\mathbb{N}, \mathbb{R}^3)$ and let $M(i)$, $i \in \mathbb{N}$, and $N \in \mathbb{N} \setminus \{0\}$ as defined in (2.6). Then

$$\begin{aligned} \forall i \in \mathbb{N} \forall \varepsilon > 0 \exists \delta > 0 \quad &: |u_k(i + N) - u_k(i)| < \delta \text{ for all } k = 1, 2, 3 \\ \implies &\|M(i + 1) - M(i)\|_2 < \varepsilon. \end{aligned} \quad (3.8)$$

Proof

Let $i \in \mathbb{N}$ and define $\Delta(i) = M(i + 1) - M(i)$. We take $|u_k(i + N) - u_k(i)| < \delta$ for some $\delta > 0$ and we prove $\|\Delta\|_2 \rightarrow 0$ ($\delta \rightarrow 0$).

We have

$$|\Delta(i)_{k,l}| \leq \delta(|u_k(i)| + |u_l(i)|) + \delta^2 \leq 2\delta \|\underline{u}\|_\infty + \delta^2 \rightarrow 0 \quad (\delta \rightarrow 0).$$

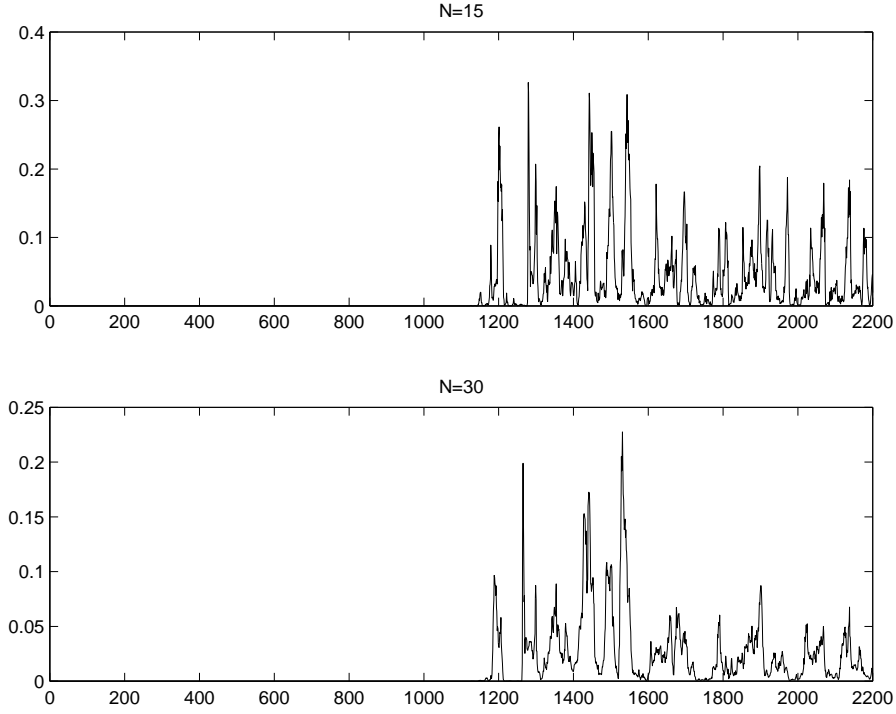


Figure 2: κ for $N = 15$ and $N = 30$

Then by Gershgorin's theorem [9]

$$\|\Delta(i)\|_2 = \rho(\Delta(i)) \leq \max_k \sum_{l=1}^3 |\Delta(i)_{k,l}| \rightarrow 0 \quad (\delta \rightarrow 0).$$

□

With this lemma we have proved that small changes in \underline{u} will lead to small changes in $M(i)$ and will yield therefore by the preceding corollaries small changes in the characteristic functions (2.7), (2.8), (2.9) and (2.10) in the evolution of time. Further we observe that for N -periodic signals \underline{u} the cross-power matrix remains the same for different $i \in \mathcal{I}N$. This fact has to be taken into account when choosing an appropriate window length for analysing \underline{u} .

3.3 Problems in analysing seismic data

In the sequel we will use three characteristic functions for the product as denoted in (2.12), namely the deflection angle (2.9), the degree of polarisation (2.10) and the energy ratio of the S-plane versus the total energy in the signal (3.2). These functions are the same functions, used by Cichowicz in [5]. From now on we will denote the product of these functions by κ .

When analysing a three-component signal with κ , we have to deal with several problems. First of all the window length N for the matrices $M(i)$ and for the energy ratio (3.2) has to be determined. Obviously N is related to the frequency spectrum of the signal, e.g. for analysing low frequency signals N must be chosen to be larger than for the analysis of high frequency signals. By fixing N we do not take into account that a signal can consist of a broad range of frequency contents. Moreover, the

window lengths N that are not related to the frequency contents of the signal will introduce undesired spikes in the graph of κ , which has been depicted in Figure 2. In this figure we see a seismic signal analysed by κ using two different window lengths, namely $N = 15$ and $N = 30$. The signal we used in this analysis is the three-component seismogram of a local event, i.e. an event that has taken place nearby the measurement equipment. Analysing the frequency spectrum will yield appropriate window lengths N in the range from 25 to 40 samples. Due to an incorrect window length we see in the upper graph a lot of local maxima that are not related to any specific features in the signal.

Another problem we have to deal with is the following. Ideally κ should attain its global maximum at i_S . Moreover, this maximal value should be close to 1. However in practice a seismic signal does not only consist of an P-phase and an S-phase, but also other waves, which we did not consider here, appear in the seismogram. Furthermore, the signal is generally measured with both background noise and signal generated noise. Due to these facts κ will generally not reach a value close to 1 at i_S and even the maximum of κ at i_S can turn out to be a local maximum instead of a global maximum. Therefore we take a threshold value that has to be reached by κ at the S-phase arrival time i_S . This situation is comparable to what we did in (2.4) for detecting the P-phase arrival time. In practice it turns out that choosing this threshold value is very difficult. To illustrate this phenomenon we have another look at Figure 2. In both pictures we notice that the global maximum of κ is much less than its ideal value 1. Furthermore, we see that in the second picture κ attains its global maximum at $i = 1540$ while $i_S = 1290$. From this example it is clear that a threshold value should be chosen very carefully.

4. THE DISCRETE WAVELET TRANSFORM

We introduce the discrete wavelet transform (DWT) as a signal processing tool. By means of filters associated with the DWT we are able to decompose seismic signals \underline{u} in time at different scales. With this decomposition we expect that it is possible to separate the S-phase from the P-phase and probably also from other phases in the signal.

4.1 Multiresolution Analysis

Before introducing the concept of a multiresolution analysis (MRA) in $L^2(\mathbb{R})$, we define two orthonormal operators on $L^2(\mathbb{R})$.

Definition 4.1

For all $f \in L^2(\mathbb{R})$ the dilation operator D_2 on $L^2(\mathbb{R})$ is given by

$$(D_2 f)(x) = \sqrt{2}f(2x), \quad \forall f \in L^2(\mathbb{R})$$

and the shift operator T on $L^2(\mathbb{R})$ is given by

$$(Tf)(x) = f(x - 1), \quad \forall f \in L^2(\mathbb{R}).$$

With these two operators we come to the definition of an MRA, following [7, 8, 14].

Definition 4.2

An MRA in $L^2(\mathbb{R})$ is an increasing sequence of closed subspaces V_j , $j \in \mathbb{Z}$, in $L^2(\mathbb{R})$,

$$\cdots \subset V_{-2} \subset V_{-1} \subset V_0 \subset V_1 \subset V_2 \cdots,$$

such that

1. $\bigcup_{j \in \mathbb{Z}} V_j$ is dense in $L^2(\mathbb{R})$,
2. $\bigcap_{j \in \mathbb{Z}} V_j = \{0\}$,
3. $f \in V_j \iff D_2 f \in V_{j+1}, \forall j \in \mathbb{Z}$,
4. $\exists_{\phi \in L^2(\mathbb{R})} : \{T^k \phi \mid k \in \mathbb{Z}\}$ is a Riesz basis for V_0 .

We recall that a Riesz basis in a Hilbert space H is a sequence of vectors x_k with closed linear span equal to H and with constants $0 < m \leq M$ such that

$$m \|\alpha\|_2 \leq \left\| \sum_{k \in \mathbb{Z}} \alpha(k) x_k \right\|_H \leq M \|\alpha\|_2, \quad (4.1)$$

for all $\alpha \in l^2(\mathbb{Z})$. Besides this definition other characterizations of a Riesz basis can be found in [19]. Since an orthonormal basis is a Riesz basis with bounds $m = M = 1$, which can be proved rather easily, an MRA in which there exists a function $\phi \in L^2(\mathbb{R})$ such that $\{T^k \phi \mid k \in \mathbb{Z}\}$ is an orthonormal basis for V_0 is a special case of an MRA as defined in Definition 4.2.

By the definition of an MRA a Riesz basis for V_j , with $j \in \mathbb{Z}$ fixed, is given by $\{D_2^j T^k \phi \mid k \in \mathbb{Z}\}$, once such function ϕ has been found. Such function ϕ is called scaling function. Since D_2 is an orthonormal operator on $L^2(\mathbb{R})$ and V_0 is invariant under the action of D_2 , the collection $\{D_2 T^k \phi \mid k \in \mathbb{Z}\}$ is a Riesz basis for V_1 . As we also have $\phi \in V_1$, we get

$$\phi = \sum_{k \in \mathbb{Z}} h(k) D_2 T^k \phi, \quad (4.2)$$

for some $h \in l^1(\mathbb{Z})$. This relation is called the scale relation and we will refer to h as the scale sequence.

We consider again the inclusion $V_0 \subset V_1$. Obviously we can define a subspaces W_0 such that $W_0 = V_1 \setminus V_0$. To come to a unique definition of W_0 , we take

$$W_0 = V_1 \cap V_0^\perp.$$

Using the invariance of the subspaces V_j under the action of the orthonormal operator D_2 we arrive in a natural way at the definition of the closed subspaces $W_j \subset L^2(\mathbb{R})$ by putting

$$W_j = V_{j+1} \cap V_j^\perp.$$

Recursively repeating the orthonormal decomposition of some V_J into V_{J-1} and W_{J-1} yields

$$\begin{aligned} V_J &= V_{J-1} \oplus W_{J-1} = V_{J-2} \oplus W_{J-2} \oplus W_{J-1} \\ &= \dots = V_{-J} \oplus \left(\bigoplus_{j=-J}^{J-1} W_j \right). \end{aligned}$$

Taking $J \rightarrow \infty$ and applying Condition 1 and 2 in Definition 4.2 leads to

$$\bigoplus_{j \in \mathbb{Z}} W_j = L^2(\mathbb{R}).$$

From this result, the question arises whether it is possible to find a function $\psi \in V_1$, such that $\{T^k\psi \mid k \in \mathbb{Z}\}$ is a Riesz basis for W_0 . Then $\{D^j T^k \psi \mid k \in \mathbb{Z}\}$ is a Riesz basis for W_j , $j \in \mathbb{Z}$. Since the subspaces W_j are chosen to be mutually orthogonal, we then have a Riesz basis in $L^2(\mathbb{R})$ given by

$$\{D^j T^k \psi \mid j, k \in \mathbb{Z}\}.$$

The function ψ will be called a wavelet and $\{D^j T^k \psi \mid k \in \mathbb{Z}\}$ a wavelet basis for W_j , for fixed $j \in \mathbb{Z}$. So, using these wavelet bases we are able to decompose any $f \in L^2(\mathbb{R})$ into functions at several scales. To get some insight in this decomposition, we can compare it with the Fourier series of f , where f is decomposed into trigonometric functions with different periods. However by using trigonometric functions we loose the localisation of certain features in the signal. By choosing an appropriate wavelet function ψ localisation is preserved.

Since the wavelet function ψ should be in V_1 , there exists also a scale relation for ψ

$$\psi = \sum_{k \in \mathbb{Z}} g(k) D_2 T^k \phi, \quad (4.3)$$

for some $g \in l^2(\mathbb{Z})$, the scaling sequence for ψ . We will also require $g \in l^1(\mathbb{Z})$. By this relation, the problem of finding ψ can be substituted by the problem of finding g if ϕ , and therefore also h , is known. For $\{T^k \phi \mid j, k \in \mathbb{Z}\}$ an orthonormal basis for V_0 a well-known choice for g is given by

$$g(k) = (-1)^k h(1 - k). \quad (4.4)$$

The derivation of this relation, see e.g. [8], uses the orthonormal decomposition of V_1 into V_0 and W_0 . In [16] we also discussed the choice of g in terms of h in the case that $\{T^k \phi \mid j, k \in \mathbb{Z}\}$ is a Riesz basis for V_0 .

We can also represent an MRA of $L^2(\mathbb{R})$ by means of projections on $L^2(\mathbb{R})$ in the following way. Corresponding to all subspaces V_j we introduce projections $P_j : L^2(\mathbb{R}) \rightarrow V_j$ given by

$$P_j f = \sum_{k \in \mathbb{Z}} (f, \tilde{x}_{j,k})_2 D_2^j T^k \phi,$$

with $\tilde{x}_{j,k} \in L^2(\mathbb{R})$ such that $(\tilde{x}_{j,k}, D_2^j T^n \phi) = \delta_{k,n}$. The sequence $\tilde{x}_{j,k}$ is not uniquely determined for a given ϕ , however by requiring $\tilde{x}_{j,k} \in V_j$, $\tilde{x}_{j,k}$ is uniquely determined and is given by

$$\tilde{x}_{j,k} = D^j T^k \left(\sum_{n \in \mathbb{Z}} \tilde{\tau}_\phi(n) T^n \phi \right),$$

where $\tilde{\tau}_\phi$ is given by the relation

$$\tau_\phi * \tilde{\tau}_\phi = e_0.$$

Here $\tau_\phi(n) = (\phi, T^n \phi)$ is a sequence in $l^2(\mathbb{Z})$ that yields a boundedly invertible convolution operator on $l^2(\mathbb{Z})$, see [16], and $\{e_n\}$ denotes the standard orthonormal basis in $l^2(\mathbb{Z})$. With this choice for the dual sequences $\tilde{x}_{j,k}$, $j \in \mathbb{Z}$, $\tilde{x}_{j,k}$ can be written as $D_2^j T^k \tilde{\phi}$, with $\{T^k \tilde{\phi} \mid k \in \mathbb{Z}\}$ a Riesz basis for V_0 . Since $\tilde{\phi} \in V_0$ there also exists a scaling relation

$$\tilde{\phi} = \sum_{k \in \mathbb{Z}} h_0(k) D_2 T^k \phi, \quad (4.5)$$

for some $h_0 \in l^1(\mathbb{Z})$.

We may verify that $\tilde{\tau} = \tau = e_0$ if $\{T^k\phi | k \in \mathbb{Z}\}$ is an orthonormal basis for V_0 . In this case P_j are orthonormal projections from $L^2(\mathbb{R})$ onto V_j given by

$$P_j f = \sum_{k \in \mathbb{Z}} (f, D_2^j T^k \phi)_2 D_2^j T^k \phi. \quad (4.6)$$

Corresponding to the subspaces W_j we can introduce the projections $Q_j : L^2(\mathbb{R}) \rightarrow W_j$ by

$$Q_j = P_{j+1} - P_j. \quad (4.7)$$

These projections can also be expressed in terms of the wavelet $\psi \in W_0$. In the same manner as we did before for P_j we can write

$$Q_j f = \sum_{k \in \mathbb{Z}} (f, \tilde{y}_{j,k})_2 D_2^j T^k \psi,$$

with $\tilde{y}_{j,k} \in L^2(\mathbb{R})$ such that $(\tilde{y}_{j,k}, D_2^j T^n \psi) = \delta_{k,n}$. By requiring $\tilde{y}_{j,k} \in W_j$, analogous to $\tilde{x}_{j,k}$, the sequence $\tilde{y}_{j,k}$ is uniquely determined and is given by

$$\tilde{y}_{j,k} = D^j T^k \left(\sum_{n \in \mathbb{Z}} \tilde{\tau}_\psi(n) T^n \psi \right),$$

where $\tilde{\tau}_\psi$ is given by

$$\tau_\psi * \tilde{\tau}_\psi = e_0$$

and with $\tau_\psi(n) = (\psi, T^n \psi)$. For τ_ψ the same properties hold as for τ_ϕ . Now we write $\tilde{y}_{j,k}$ as $D_2^j T^k \tilde{\psi}$, where $\{T^k \tilde{\psi} | k \in \mathbb{Z}\}$ is a Riesz basis for W_0 . Also for $\tilde{\psi} \in W_0$ there exists a scaling relation

$$\tilde{\psi} = \sum_{k \in \mathbb{Z}} g_0(k) D_2 T^k \phi, \quad (4.8)$$

for some $g_0 \in l^1(\mathbb{Z})$.

As before we see, that for $\{T^k \psi | k \in \mathbb{Z}\}$ an orthonormal basis for W_0 the projections Q_j become orthonormal projections from $L^2(\mathbb{R})$ onto W_j given by

$$Q_j f = \sum_{k \in \mathbb{Z}} (f, D_2^j T^k \psi)_2 D_2^j T^k \psi. \quad (4.9)$$

4.2 Decomposition and reconstruction

We discuss fast methods to decompose $P_{J+1}f$ into $P_J f$ and $Q_J f$ and to recover $P_{J+1}f$ from $P_J f$ and $Q_J f$ by means convolution products.

4.2.1 The decomposition algorithm: We assume $P_{J+1}f \in V_{J+1}$ is known for a certain $J \in \mathbb{Z}$. So there exists a sequence $c_{J+1} \in l^2(\mathbb{Z})$ such that

$$P_{J+1}f = \sum_{k \in \mathbb{Z}} c_{J+1}(k) D_2^{J+1} T^k \phi.$$

Moreover, the sequence c_{J+1} is given by $c_{J+1}(k) = (D_2^{J+1}T^k\tilde{\phi}, f)_2$. Following (4.7), $P_{J+1}f$ can be decomposed into a function in V_J given by

$$P_Jf = \sum_{k \in \mathbb{Z}} c_J(k) D_2^J T^k \phi,$$

and into a function in W_J given by

$$Q_Jf = \sum_{k \in \mathbb{Z}} d_J(k) D_2^J T^k \psi.$$

The sequences c_J and d_J can be calculated respectively by $c_J(k) = (D_2^J T^k \tilde{\phi}, f)_2$ and $d_J(k) = (D_2^J T^k \tilde{\psi}, f)_2$. However there is also a fast method to derive c_J and d_J from the known sequence c_{J+1} without calculating inner products with f .

To get an expression for the coefficients c_J in terms of c_{J+1} we derive $(P_Jf, D_2^J T^k \tilde{\phi})_2 = c_J(k)$ and $(Q_Jf, D_2^J T^k \tilde{\phi})_2 = 0$. These results yield in combination with (4.5) and (4.7)

$$\begin{aligned} c_J(k) &= (P_{J+1}f, D_2^J T^k \tilde{\phi})_2 = \sum_{n \in \mathbb{Z}} h_0(n) (P_{J+1}f, D_2^{J+1} T^{2k+n} \phi)_2 \\ &= \sum_{m, n \in \mathbb{Z}} c_{J+1}(m) h_0(n) (D_2^{J+1} T^m \phi, D_2^{J+1} T^{2k+n} \phi)_2 \\ &= \sum_{m, n \in \mathbb{Z}} c_{J+1}(m) h_0(n) \tau_\phi(2k + n - m) \\ &= \sum_{n \in \mathbb{Z}} (c_{J+1} * \tau_\phi)(2k + n) h_0(n) = (c_{J+1} * \tau_\phi * \check{h}_0)(2k), \end{aligned} \quad (4.10)$$

with $\check{h}_0(n) = h_0(-n)$. It follows from (4.10) that the coefficients c_J can be computed by taking the even entries of the convolution product of c_{J+1} with τ_ϕ and \check{h}_0 .

For the coefficients d_J we get a similar relation by taking the inner product of $D_2^J T^k \tilde{\phi}$ with both the left-hand side and the right-hand side of equation (4.7). This derivation, that involves (4.8), yields

$$d_J(k) = (c_{J+1} * \tau_\phi * \check{g}_0)(2k). \quad (4.11)$$

Since our derivations hold for any $J \in \mathbb{Z}$ we can rather easily compute all coefficients c_j and d_j that appear in P_jf and Q_jf for $j < J$ once c_J is known.

We observe that in the case ϕ and ψ generate orthonormal systems by means of their translates, relation (4.10) and (4.11) will simplify to

$$\begin{aligned} c_J(k) &= (c_{J+1} * \check{h})(2k), \\ d_J(k) &= (c_{J+1} * \check{g})(2k). \end{aligned}$$

4.2.2 The reconstruction algorithm: Once we have computed a decomposition of $P_{J+1}f$ into P_Jf and Q_Jf by means of the coefficients c_J and d_J , we can also recover c_{J+1} out of c_J and d_J in an efficient way. In order to come to such a reconstruction formula we will represent P_Jf and Q_Jf in terms of $D_2^{J+1}T^k\phi$, $k \in \mathbb{Z}$, the basis functions of V_{J+1} .

Since $P_J f \in V_J$ and $V_J \subset V_{J+1}$, there exists an $\alpha_J \in l^2(\mathbb{Z})$ such that

$$P_J f = \sum_{k \in \mathbb{Z}} \alpha_J(k) D^{J+1} T^k \phi.$$

Taking the inner product with $D^{J+1} T^k \tilde{\phi}$ in both the left hand side and the right hand side of this equation yields in combination with (4.2)

$$\begin{aligned} \alpha_J(k) &= (P_J f, D^{J+1} T^k \tilde{\phi})_2 = \sum_{n \in \mathbb{Z}} c_J(n) (D_2^J T^n \phi, D^{J+1} T^k \tilde{\phi})_2 \\ &= \sum_{m, n \in \mathbb{Z}} c_J(n) h(m) (D_2^{J+1} T^{m+2n} \phi, D^{J+1} T^k \tilde{\phi})_2 \\ &= \sum_{n \in \mathbb{Z}} c_J(n) h(k - 2n) = (\bar{c}_J * h)(k), \end{aligned} \tag{4.12}$$

with \bar{c}_J such that $\bar{c}_J(2n) = c_J(n)$ and $\bar{c}_J(2n + 1) = 0$ for all $n \in \mathbb{Z}$.

In the same manner we can search for a sequence $\beta_J \in l^2(\mathbb{Z})$ such that

$$Q_J f = \sum_{k \in \mathbb{Z}} \beta_J(k) D^{J+1} T^k \phi.$$

Following the derivation of (4.12) we get for β_J the expression

$$\beta_J = \bar{d}_J * g. \tag{4.13}$$

To reconstruct c_{J+1} from c_J and d_J we use (4.7) in combination with the preceding results. This leads to the reconstruction formula

$$c_{J+1} = \bar{c}_J * h + \bar{d}_J * g. \tag{4.14}$$

Again we observe that our derivations hold for any $J \in \mathbb{Z}$, so that we can rather easily recover all coefficients c_{j+1} from c_j and d_j for any $j \in \mathbb{Z}$.

4.3 The filter bank approach

The discussed decomposition and reconstruction methods can also be interpreted in terms of filter banks, see [12, 17]. For this interpretation we introduce two operators from signal processing.

The first operator we introduce is the downsampling operator given by

$$((\downarrow t)\alpha)(k) = \alpha(kt),$$

for $t \in \mathbb{N}$ and for all $\alpha \in l^2(\mathbb{Z})$. For this linear operator we have $\|(\downarrow t)\|_2 \leq 1$. With the downsampling operator we can rewrite the decomposition formulas (4.10) and (4.11) as

$$\begin{aligned} c_J &= (\downarrow 2)(c_{J+1} * \tau_\phi * \check{h}_0) \\ d_J &= (\downarrow 2)(c_{J+1} * \tau_\phi * \check{g}_0) \end{aligned}$$

The other linear operator we discuss is the upsampling operator given by

$$((\uparrow t)\alpha)(k) = \begin{cases} \alpha(k/t), & \text{if } k \bmod t = 0, \\ 0, & \text{otherwise,} \end{cases}$$

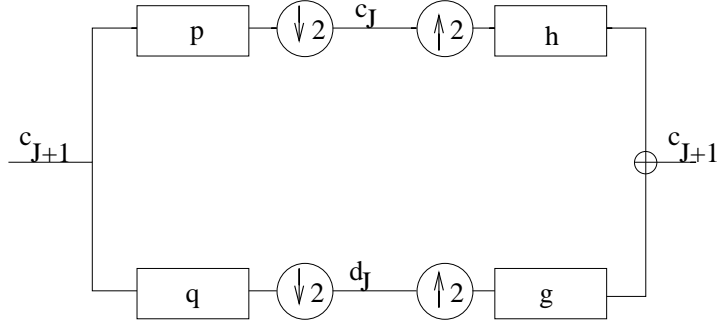


Figure 3: Decomposition/reconstruction filter bank

for $t \in \mathbb{N}$ and for all $\alpha \in l^2(\mathbb{Z})$. We observe that $(\uparrow t)$ is an isometry for all $t \in \mathbb{N}$ and that $(\uparrow t)^* = (\downarrow t)$. Using the upsampling operator we can rewrite the reconstruction formulas (4.12) and (4.13) as

$$\begin{aligned}\alpha_J &= h * ((\uparrow 2)c_J) \\ \beta_J &= g * ((\uparrow 2)d_J)\end{aligned}$$

The method for decomposing and reconstructing c_{J+1} can be seen as the analysis and synthesis part of a two channel filter bank as has been depicted in Figure 3. In the first channel of the analysis part c_{J+1} is convolved with the filter $p = \tau_\phi * \check{h}_0$ and afterwards downsampled by 2 to obtain c_J . In the second channel the filter $q = \tau_\phi * \check{g}_0$ is used instead of p to obtain d_J . In the synthesis part c_J and d_J are first upsampled by 2 and afterwards filtered by h and g respectively. At the end of the synthesis part we get α_J as output for the first channel and β_J in the second channel.

The discussed filter bank can also be used to decompose c_J into c_{J-1} and d_{J-1} . For this we take c_J as input instead of c_{J+1} . After filtering and down sampling we get the sequences c_{J-1} and d_{J-1} respectively in the first and second channel of the filter bank. In the synthesis part c_{J-1} and d_{J-1} are upsampled and filtered by h and g respectively to obtain α_{J-1} in the first channel and β_{J-1} in the second channel of the filter bank. Using this recursive process we get the following expressions for c_j , d_j , α_j and β_j for $j \leq J$.

$$c_j = ((\downarrow 2)\mathcal{C}_p)^{J+1-j}c_{J+1}, \quad (4.15)$$

$$d_j = (\downarrow 2)\mathcal{C}_q((\downarrow 2)\mathcal{C}_p)^{J-j}c_{J+1}, \quad (4.16)$$

$$\alpha_j = (\mathcal{C}_h(\uparrow 2))^{J+1-j}((\downarrow 2)\mathcal{C}_p)^{J+1-j}c_{J+1}, \quad (4.17)$$

$$\beta_j = (\mathcal{C}_h(\uparrow 2))^{J-j}\mathcal{C}_g(\uparrow 2)(\downarrow 2)\mathcal{C}_q((\downarrow 2)\mathcal{C}_p)^{J-j}c_{J+1}, \quad (4.18)$$

with \mathcal{C}_x denoting the convolution with x . Since we required $g, h \in l^1(\mathbb{Z})$, all sequences above are in $l^2(\mathbb{Z})$. Note that in the case $\{T^k\phi \mid k \in \mathbb{Z}\}$ and $\{T^k\psi \mid k \in \mathbb{Z}\}$ are orthonormal systems, we have $p = \check{h}$ and $q = \check{g}$. In the next section we will see how these formulas are used in our algorithm to represent time-discrete seismograms at several scales.

5. THE DWT FOR SEISMIC ANALYSIS

We discuss how the introduced DWT can be used in combination with the characteristic functions of Section 2 to analyse three-component seismograms. Since the seismic data are discrete-time functions we first consider the relation between the presented DWT and the analysis of sampled signals at scales. After that we discuss how the characteristic function κ , as denoted in (2.12), can be used in an efficient way to analyse seismic data at several scales.

5.1 Analysing discrete-time signals with the DWT

In Section 4 we introduced the DWT for functions $f \in L^2(\mathbb{R})$. In our problem we are dealing with time-discrete signals in $l^2(\mathbb{Z})$ obtained by sampling f . In order to analyse these signals in $l^2(\mathbb{Z})$ at several scales we introduce a decomposition at scales for sequences in $l^2(\mathbb{Z})$.

Definition 5.1

Let $s \in l^2(\mathbb{Z})$. Then the l^2 -DWT of s at scale m is defined as

$$s^{(m)} = (C_h(\uparrow 2))^{m-1} C_g(\uparrow 2)(\downarrow 2) C_q((\downarrow 2) C_p)^{m-1} s, \quad (5.1)$$

for $m \in \mathbb{Z}$. Here h and g denote the scale sequences as given in (4.2) and (4.3) and p and q are given respectively by $p = \tau_\phi * \check{h}_0$ and $q = \tau_\phi * \check{g}_0$, cf. their definitions in Section 4.3.

We observe that (4.18) can also be seen as the l^2 -DWT of c_{J+1} at scale $J+1-j$, i.e. $\beta_j = c_J^{(J+1-j)}$.

Besides the l^2 -DWT of a sequence $s \in l^2(\mathbb{Z})$, we also define its approximation sequence $s_a^{(m)}$ at level m by

$$s_a^{(m)} = (C_h(\uparrow 2))^m ((\downarrow 2) C_p)^m s. \quad (5.2)$$

Note that (4.17) can also be seen as the approximation sequence of C_J at level $J+1-j$, i.e. $\alpha_j = (c_J)_a^{(J+1-j)}$.

The l^2 -DWT we proposed in this paper can also be seen as a special case of the DWT for $l^2(\mathbb{Z})$, that Cohen proposed in [7] using a dyadic dilation operator on $l^2(\mathbb{Z})$ and the shift operator R to construct an MRA for $l^2(\mathbb{Z})$. However since Cohen's approach only deals with orthonormal wavelet bases, we cannot use all results on the DWT for $l^2(\mathbb{Z})$ from [7].

Following the recursive filter bank approach of Section 4.3 we get

$$s = s_a^{(M)} + \sum_{m=1}^M s^{(m)},$$

for all $M \in \mathbb{N}$, by substituting $s = c_{J+1}$. In [4], Frazier and Kumar showed that sequences generated by an iterated filter bank approach, like $s_a^{(M)}$ and $s^{(m)}$, $m = 1, \dots, M$ are linearly independent if $\{R^{2k}g \mid k \in \mathbb{Z}\}$ and $\{R^{2k}h \mid k \in \mathbb{Z}\}$ are linearly independent sets in $l^2(\mathbb{Z})$. Moreover, they proved that if $\{R^{2k}g \mid k \in \mathbb{Z}\}$ and $\{R^{2k}h \mid k \in \mathbb{Z}\}$ are orthonormal systems in $l^2(\mathbb{Z})$, with also $(g, R^{2k}h)_2 = 0$ for all $k \in \mathbb{Z}$, then $s_a^{(M)}$ and $s^{(m)}$, $m = 1, \dots, M$, are mutually orthonormal. Here R denotes the shift operator on $l^2(\mathbb{Z})$ given by $(R\alpha)(k) = \alpha(k-1)$. As a corollary we get

$$\|s\|_2^2 = \|s_a^{(M)}\|_2^2 + \sum_{m=1}^M \|s^{(m)}\|_2^2, \quad (5.3)$$

saying that the energy in s is equal to the energy of its decomposition signals at each scale using the l^2 -DWT plus its approximation at a certain level. This result can also be related to the wavelet function ψ and its scaling function ϕ . In [16] we proved the following lemma.

Lemma 5.2

Let V_j be closed subspaces in $L^2(\mathbb{R})$ as defined in Definition 4.2, and let W_j be given by $W_j = V_{j+1} \cap V_j^\perp$. Let further $\{T^k\phi \mid k \in \mathbb{Z}\}$ and $\{T^k\psi \mid k \in \mathbb{Z}\}$ be orthonormal bases for V_0 and W_0 respectively. Then

$$\{R^{2k}g \mid k \in \mathbb{Z}\} \cup \{R^{2k}h \mid k \in \mathbb{Z}\}$$

is an orthonormal basis for $l^2(\mathbb{Z})$. The converse also holds.

Combining Lemma 5.2 with (5.3) we arrive at

Corollary 5.3

Let V_j be closed subspaces in $L^2(\mathbb{R})$ as defined in Definition 4.2, and let W_j be given by $W_j = V_{j+1} \cap V_j^\perp$. Let further $\{T^k\phi \mid k \in \mathbb{Z}\}$ and $\{T^k\psi \mid k \in \mathbb{Z}\}$ be orthonormal bases for V_0 and W_0 respectively. Then

$$\|s\|_2^2 = \|s_a^{(M)}\|_2^2 + \sum_{m=1}^M \|s^{(m)}\|_2^2$$

and conversely.

In the case $\{T^k\phi \mid k \in \mathbb{Z}\}$ and $\{T^k\psi \mid k \in \mathbb{Z}\}$ are Riesz systems (5.3) does not hold in general. To get some insight why (5.3) only holds in general for orthonormal wavelet systems, we consider the case $M = 1$. We rewrite $s^{(1)}$ and $s_a^{(1)}$ in terms of the basis sequences $R^{2k}g$ and $R^{2k}h$, $k \in \mathbb{Z}$, namely

$$\begin{aligned} s^{(1)} &= \sum_{k \in \mathbb{Z}} (s, R^{2k}\check{q}) R^{2k}g, \\ s_a^{(1)} &= \sum_{k \in \mathbb{Z}} (s, R^{2k}\check{p}) R^{2k}h. \end{aligned}$$

Since $s = s^{(1)} + s_a^{(1)}$ we get by taking inner products

$$\|s\|_2^2 = \|s_a^{(1)}\|_2^2 + \|s^{(1)}\|_2^2 + 2 \sum_{k,l \in \mathbb{Z}} (s, R^{2k}\check{p})_2 (s, R^{2l}\check{q})_2 (g, R^{2(l-k)}h)_2.$$

For (5.3) to hold for any $s \in l^2(\mathbb{Z})$, the inner product $(g, R^{2n}h)_2$ should vanish for any $n \in \mathbb{Z}$. However in [16] we showed that in this case $(\tau_\phi * g, R^{2n}h)_2 = 0$ for all $n \in \mathbb{Z}$.

From (5.3) we derive straightforward $\|s_a^{(M)}\|_2 \geq \|s_a^{(M+1)}\|_2 \geq 0$. This yields that $\lim_{m \rightarrow \infty} \|s^{(m)}\|_2$ exists. In the following theorem we show that $\lim_{m \rightarrow \infty} \|s^{(m)}\|_2 = 0$, following [7].

Theorem 5.4

Let $s \in l^2(\mathbb{Z})$ and $h \in l^1(\mathbb{Z})$, such that $\{R^{2k}h \mid k \in \mathbb{Z}\}$ is an orthonormal system in $l^2(\mathbb{Z})$. Let further $s_a^{(m)}$ be given by

$$s_a^{(m)} = (\mathcal{C}_h(\uparrow 2))^m ((\downarrow 2)\mathcal{C}_h)^m s$$

for $m \in \mathbb{N}$. Then

$$\|s_a^{(m)}\|_2 \rightarrow 0 \quad (m \rightarrow \infty). \quad (5.4)$$

Proof

We write $\mathcal{A} = (\downarrow 2)\mathcal{C}_{\tilde{h}}$. Then $\mathcal{A}^* = \mathcal{C}_h(\uparrow 2)$. So $s_a^{(m)} = (\mathcal{A}^*)^m \mathcal{A}^m s$. By the orthonormality of $R^{2k}h$, $k \in \mathbb{Z}$, we also get $\mathcal{A}\mathcal{A}^* = \mathcal{I}$ and thus

$$\|s_a^{(m)}\|_2^2 = ((\mathcal{A}^*)^m \mathcal{A}^m s, (\mathcal{A}^*)^m \mathcal{A}^m s)_2 = (\mathcal{A}^m s, \mathcal{A}^m s)_2 = \|\mathcal{A}^m s\|_2^2.$$

Cohen proved in [7], that $\|\mathcal{A}^m \delta_k\|_2$ tends to zero as $m \rightarrow \infty$, with $\{\delta_k \mid k \in \mathbb{Z}\}$ the standard orthonormal basis in $l^2(\mathbb{Z})$. So $\|s_a^{(m)}\|_2 \rightarrow 0$ ($m \rightarrow \infty$) for any $s \in l_0^2(\mathbb{Z}) = \text{span}\{\delta_k \mid k \in \mathbb{Z}\}$. Since $l_0^2(\mathbb{Z})$ is dense in $l^2(\mathbb{Z})$ the result holds for any $s \in l^2(\mathbb{Z})$.
□

The following corollary will be of interest when implementing our algorithm, since it says that if (5.3) holds, the energy will tend to zero at large scaling levels.

Corollary 5.5

Let $s \in l^2(\mathbb{Z})$ and $g, h \in l^1(\mathbb{Z})$, such that

$$\{R^{2k}g \mid k \in \mathbb{Z}\} \cup \{R^{2k}h \mid k \in \mathbb{Z}\}$$

is an orthonormal basis of $l^2(\mathbb{Z})$. Let further $s^{(m)}$ be given by

$$s^{(m)} = (\mathcal{C}_h(\uparrow 2))^{m-1} \mathcal{C}_g(\uparrow 2) (\downarrow 2) \mathcal{C}_{\tilde{g}} ((\downarrow 2) \mathcal{C}_{\tilde{h}})^{m-1} s$$

for $m \in \mathbb{N}$. Then

$$\|s^{(m)}\|_2 \rightarrow 0 \quad (m \rightarrow \infty). \quad (5.5)$$

Proof

Combining (5.3) and (5.4) yields

$$\lim_{M \rightarrow \infty} \sum_{m=1}^M \|s^{(m)}\|_2^2 = \|s\|_2^2.$$

So $\|s^{(m)}\|_2$ will tend to zero for $m \rightarrow \infty$.
□

5.2 Characteristic functions and the DWT

The idea to detect S-phase arrival times using the DWT is as follows. By making a decomposition of the three-component seismic signal \underline{u} into signals at several scales, it is possible to separate the S-phase from other phases, that appear at other scales. This decomposition is made by taking the l^2 -DWT of each component of \underline{u} after rotation into longitudinal and transverse directions, i.e. $\underline{u}^{(m)} = u_L^{(m)} \underline{v}_1(i_P) + u_Q^{(m)} \underline{v}_2(i_P) + u_T^{(m)} \underline{v}_3(i_P)$.

It is already known that the frequency spectrum of the P-phase appears at higher frequencies than the S-phase. Therefore in general the S-phase will appear at higher scales than the P-phase. Our aim is now to find the set J_S , denoting the collection of scaling levels at which most of the S-phase appears. To make this choice more explicit, we compute the energy in the l^2 -DWT of a segment \underline{w} of \underline{u} for each $m \in \mathbb{N}$. This segment is given by

$$\underline{w}(i) = \underline{u}(i + i_P), \quad i = 0, \dots, T - 1,$$

with T such that $T/f_s = 25$, for the sampling frequency f_s . So we compute the l^2 -DWT of a segment of 25 seconds after the P-phase arrival time. For local events the S-phase arrival time can be noticed within this time period. For non-local events the difference $i_S - i_P$ will be larger than 25 seconds in general. Another difference between local and non-local events is the frequency behaviour of the phases. Both the P-phase and the S-phase for local events are high frequency signals compared to the P-phase and S-phase of non-local events. Therefore most of the energy of \underline{w} is found at low scaling levels for local events and at higher scales for non-local events.

For orthonormal scaling and wavelet functions we compute the energy distribution of \underline{w} over all scales

$$E(m) = \sum_{n=1}^3 \|w_n^{(m)}\|_2^2 / \sum_{n=1}^3 \|w_n\|_2^2. \quad (5.6)$$

Following Corollary 5.3 and Theorem 5.4 we get $\sum_{m=1}^{\infty} E(m) = 1$. Following Corollary 5.5, the energy distribution will tend to zero once a certain scaling level has been reached. At that moment we can stop computing the l^2 -DWT for higher scales. In Figure 4 the energy distribution E as a function of the scaling level has been depicted for a local event. In this analysis the Daubechies-4 (Db4) wavelet filter and its corresponding scaling filter have been used, see [8]. We observe that almost all energy in \underline{w} is contained in the first five scales. Note that other wavelet filters yield other energy distributions for the same seismic signal. Therefore we can only indicate the scales at which local and non-local events by experiments with a chosen wavelet filter.

To analyse the energy distribution we consider the scale m_{max} for which the maximum of E is attained. In case of a local event this maximum will be related to the S-phase, which contains most of the energy in \underline{w} . For non-local event m_{max} can be related to the P-phase, since the S-phase is not contained in \underline{w} . Let us now assume that we know that most of the energy of local events can be found generally at the first m_{loc} scales. Then in general an event can be characterised as a local event if $m_{max} \leq m_{loc}$ and as a non-local event if $m_{max} > m_{loc}$.

Knowing which event we are dealing with, we choose $J_S = \{m_{max}\}$ for local events and $J_S = \{m_{max} + 1, m_{max} + 2\}$. The latter choice is justified by observing that the S-phase will appear at higher scales than the P-phase, that is related to m_{max} . Resuming, we take

$$J_S = \begin{cases} m_{max}, & \text{if } m_{max} \leq m_{loc}, \\ \{m_{max} + 1, m_{max} + 2\}, & \text{if } m_{max} > m_{loc}. \end{cases} \quad (5.7)$$

Once J_S has been determined a product of characteristic functions (2.12), denoted by κ , can be applied to $\underline{u}^{(m)}$ for $m \in J_S$ and the S-phase arrival time is then be given by.

$$\prod_{m \in J_S} (\kappa(\underline{u}^{(m)}))^2(i) = \max_{n \in \mathbb{N}} \prod_{m \in J_S} (\kappa(\underline{u}^{(m)}))^2(n) \iff i = i_S. \quad (5.8)$$

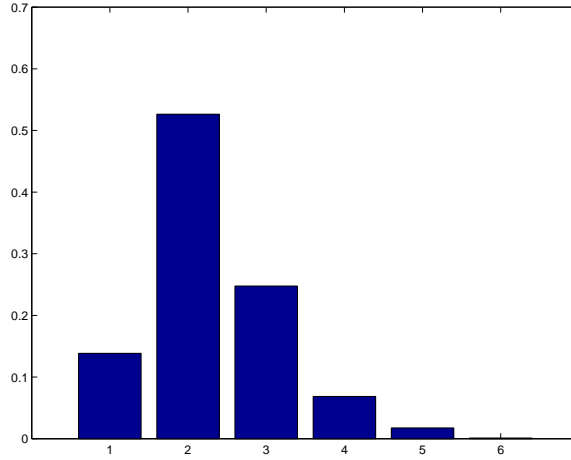


Figure 4: Energy distribution local event

In practice we also have to work with a threshold value instead of the maximum in (5.8) that has to be reached to declare i_S . However by separating the S-phase from the P-phase it will be less difficult to find an appropriate threshold value than in the situation we discussed in Section 3.3.

Another problem discussed in Section 3.3 is the window length N which was not related to the frequency behaviour of \underline{u} in traditional methods [5]. Once we are using a decomposition of \underline{u} at several scales, we can use window lengths $N(m)$, i.e. a monotone ascending function of the scaling level m . So the characteristic function κ to analyse $\underline{u}^{(m)}$ will use the window length $N(m)$ in its definition. In this manner the window length is adapted automatically to the frequency behaviour of the signal. In the next section we will discuss the choice we made for $N(m)$.

5.3 A set-up for the DWT analysis

In the previous section we discussed our approach to analyse three-component seismic data using characteristic functions and the DWT. Here we will present some choices we made for the parameters in our algorithm after testing the method on seismic data. Note that also other choices for the parameters can be made as long as they fit in the mathematical framework of our algorithm.

5.3.1 Wavelet filters: The first parameter we discuss is the wavelet function and the associated wavelet filters. We want the wavelet function to match with the seismic data we are dealing with, so that the dilated and translated wavelets generate a good approximation of the several phases at the particular scaling levels. Another desirable property for the wavelet and scaling function is orthonormality. With orthonormal wavelets, the filters p and q simplify to \check{h} and \check{g} respectively. Furthermore (5.3) holds for orthonormal wavelets and so we can use the energy distribution E to determine J_S as described in the previous section. In experiments we minimised for a set of seismic data (also synthetic data) for which i_S is known the error in the determined i_S for a collection of candidate wavelet filters. The experimental results led to the choice of the Daubechies-4 wavelet filters. These are finite impulse response filters (FIR) of length 8. So in (5.1) we only need to compute convolutions of s with sequences that contain only 8 non-zero elements.

5.3.2 S-phase scales J_S : According to (5.7) we only need to determine the scaling level m_{loc} . After having chosen the Db4 wavelet filters we computed the energy distribution (5.6) for the same set of

seismic data we used to determine an appropriate wavelet filter. For this set we found $m_{max} \leq 4$ for local events and $m_{max} > 4$ for non-local events. Therefore we take $m_{loc} = 4$ in the set-up of our algorithm.

5.3.3 Characteristic functions: Another choice we have to make before using the algorithm is which characteristic functions we want to use in (2.12). We observe that there is a natural trade-off between the robustness of (2.12) and the computing time. Every additional characteristic function could neglect artifacts that influence the determination of i_S . However every characteristic function will also contribute to the computing time of our algorithm. In our set-up we follow Cichowicz [5] by choosing three reliable characteristic functions, namely the deflection angle (2.9), the degree of polarisation (2.10) and the fraction of energy in the S-plane to the total amount of energy in the seismic signal (3.2).

5.3.4 Window length $N(m)$: In Section 5.2 we already considered a way to adapt the window length N used in the characteristic functions to the frequency behaviour of the signals. Obviously, to analyse \underline{u} at several scales $m \in \mathbb{N}$ the window length $N(m)$ has to be a monotone ascending function of the scaling level m . In order to come to an appropriate function for $N(m)$ one has to consider two facts. The signal is scaled at each level by the factor 2. Further, for low and high scaling levels $N(m)$ should not become too small or too big, since $N(m)$ is used to obtain information out of the signal in a certain neighbourhood. These considerations led to the choice

$$N(m) = \lceil 30 \cdot 2^{\max(0, (m-4)/2)} \rceil. \quad (5.9)$$

So for the first four levels we take $N = 30$ and thereafter N is multiplied by powers of $\sqrt{2}$.

5.3.5 Threshold value: To declare an S-phase arrival time we can use (5.8). However, as we discussed already in practice κ will not attain its maximum value κ_{max} at i_S . Therefore we replace κ_{max} by a threshold value $0 \leq \gamma \leq \kappa_{max}$. Since κ_{max} will vary for a set of seismic data we also want to adapt γ . This can be done by putting

$$\gamma = c \cdot \kappa_{max}. \quad (5.10)$$

Now the problem is left to choose c . First experiments with $0.6 < c < 0.9$ worked out very well, however more experiments are needed to come to an efficient value for c .

6. FIRST RESULTS

The algorithm has been implemented in *Matlab* with the set-up as discussed in Section 5.3. The implementation has been tested for a relatively small number of seismic events (15), both local and non-local, provided by the Royal Dutch Meteorological Institute (KNMI). Also Cichowicz' method [5] has been implemented in a *Matlab* code in order to compare the wavelet based method with the traditional approach.

First results look very promising. Our method failed to detect the S-phase arrival time for only one event. However also Cichowicz' method failed to detect the S-phase arrival in this case. Further, for these events it turned out that choosing the parameter c in (5.10) is easier for our method than for Cichowicz' approach. Once a good threshold value and window length had been chosen in Cichowicz' method, generally the error in i_S for our approach turned out to be slightly smaller than the error in

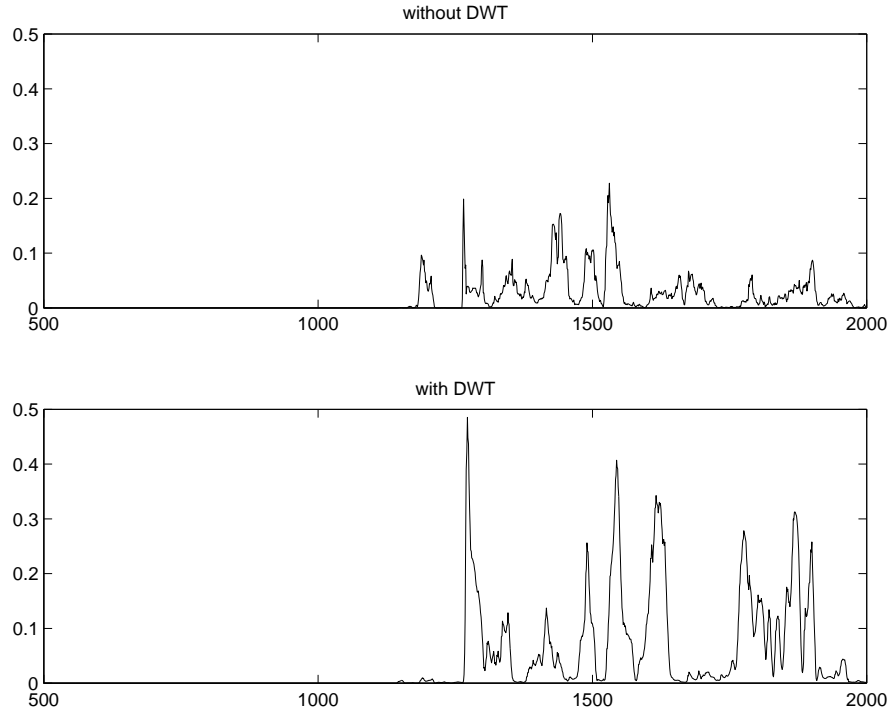


Figure 5: Characteristic function of a local event

i_S found with the traditional method [5].

To illustrate the difference of our approach compared to Cichowicz' approach, we have plotted both the characteristic function in (2.12) and the characteristic function in (5.8) for two events. For both functions we used the set-up as discussed in Section 5.3.3.

In Figure 5 we see characteristic functions of the local event which has been depicted in Figure 1. Both functions use the window length $N = 30$ and after analysing the energy distribution along scales (5.6) we got $J_S = \{2\}$. Obviously, the parameter c in (5.10) can attain almost any value between 0 and 1 in our approach, while the only possible choices for c in Cichowicz' approach is $0.55 < c < 0.95$. Note, that for automatic phase detection we have to choose a value for c before analysing a seismic event.

For this event an analyst took $i_S = 1275$. For our approach we get $i_S = 1280$ and for Cichowicz' approach we have $i_S = 1295$. With a 40 Hz sampling frequency these results differ 0.125 and 0.5 seconds respectively. We see that the results do not differ very much. However in our approach it was much easier to determine the S-phase arrival time than in the traditional approach.

In Figure 7 we see characteristic functions of the teleseismic event depicted in Figure 6. The first two functions are based on Cichowicz' method using the window lengths $N = 60$ and $N = 120$ respectively. These values for N correspond to $N(6)$ and $N(8)$, cf. (5.9). The third function uses our DWT based approach. Computing (5.6) yielded $J_S = \{7, 8\}$, and so the product of the characteristic function at scale 7 and 8 has been computed and plotted in Figure 7. We see that in the first picture the S-phase arrival will not be detected unless we put $0.96 < c \leq 1$. With a larger window $N = 120$,

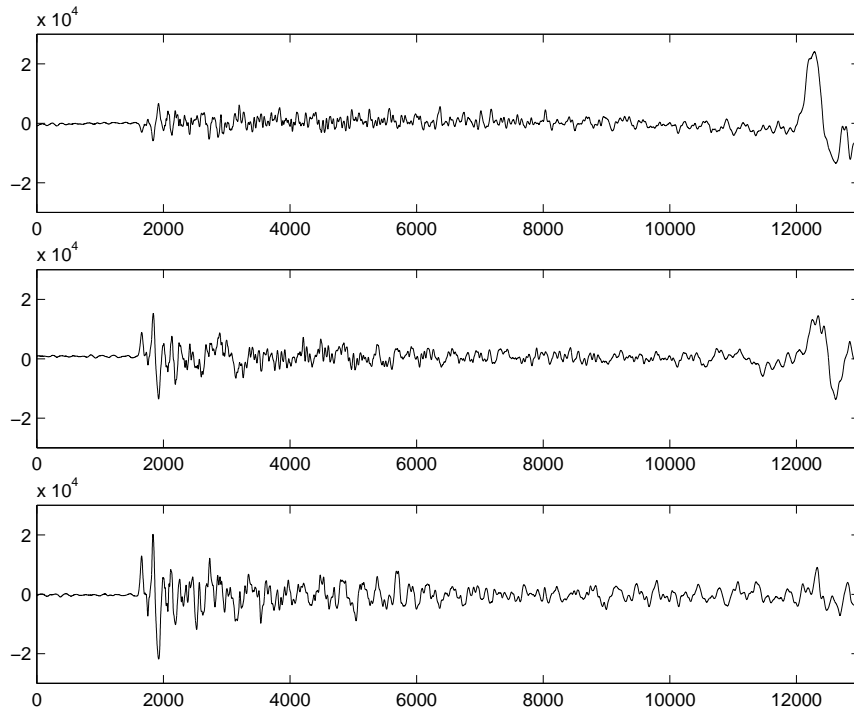


Figure 6: A teleseismic event

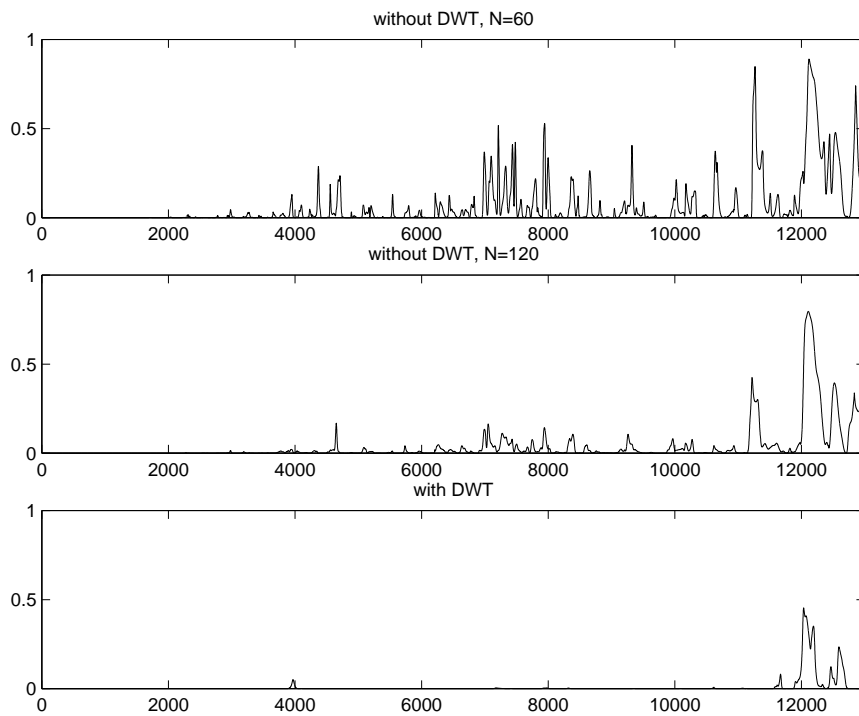


Figure 7: Characteristic function of a teleseismic event

the S-phase arrival can be detected if we take $0.55 < c \leq 1$, however in our approach we do not have to be that precise with choosing c , since for all $0.15 < c \leq 1$ the S-phase arrival will be detected.

An analyst took $i_S = 11900$ for this particular event. We get $i_S = 11960$ for our wavelet approach and $i_S = 11990$ for Cichowicz' approach. Since also this signal has been sampled at 40 Hz, these results differ 1.5 and 2.25 seconds respectively. For a teleseismic event both results are very appropriate.

At the moment our method is implemented at KNMI and tested for a huge set (300) of local events. The results of this test will be used to improve the set-up for our method.

7. CONCLUSIONS

We presented a new method to detect S-phase arrival times in three-component seismograms. This method is based on both traditional methods from seismology based on physical properties of seismic signals and on the discrete wavelet transform. This transform has been implemented using filter operations on the sampled seismograms.

First results show that this method can be very reliable for automatically detecting S-phases in a seismogram, however there are still some unsolved questions on how to pick the optimal values for the several parameters in the algorithm. We expect that further testing on huge seismic datasets will lead to better choices for the parameters involved in our algorithm.

ACKNOWLEDGEMENTS

The author likes to thank the KNMI for providing seismic data to test the algorithm. In particular he wants to thank Reinoud Sleeman and Torild Van Eck from the Seismology Department of the KNMI for their advices and suggestions in the field of seismology.

REFERENCES

1. K. Aki & P.G. Richards, *Quantitative seismology*, Freeman and Company, San Francisco, 1980.
2. K. Anant & F. Dowla, "Wavelet transform methods for phase identification in three-component seismograms", *Bull. Seism. Soc. Am.*, 87(6), 1598-1612, 1997.
3. M. Bear & U. Kradolfer, "An automatic phase picker for local and teleseismic events", *Bull. Seism. Soc. Am.*, 77(4), 1437-1445, 1987.
4. J.J. Benedetto & M.W. Frazier, *Wavelets: Mathematics and Applications*, CRC Press, London, 1994.
5. A. Cichowicz, "An automatic S-phase picker", *Bull. Seism. Soc. Am.*, 83(1), 180-189, 1993.
6. A. Cichowicz, R.W.E. Green & A. van Zyl Brink, "Coda polarization properties of high-frequency microseismic events", *Bull. Seism. Soc. Am.*, 78, 1297-1318, 1988.
7. A. Cohen & R. Ryan, *Wavelets and multiscale signal processing*, Chapman & Hall, London, 1995.
8. I. Daubechies, *Ten lectures on wavelets*, SIAM, Philadelphia, 1992.
9. J. Franklin, *Matrix Theory*, Prentice Hall, Englewood Cliffs, 1968.

10. D.C. Jepsen & B.L.N. Kennett, "Three-component analysis of regional seismograms", *Bull. Seism. Soc. Am.*, 80, 2032-2052, 1990.
11. A. Jurkevics, "Polarization analysis of three-component array data", *Bull. Seism. Soc. Am.*, 78, 1725-1743, 1988.
12. C. Herley & M. Vetterli, "Wavelets and filter banks: theory and design", *IEEE Trans. Sig. Proc.*, 40, 2207-2232, 1992.
13. E. Kanasevich, *Time sequence analysis in geophysics*, The University of Alberta Press, Alberta, 1981.
14. S. Mallat, "A theory for multiresolution signal decomposition: the wavelet representation", *IEEE Trans. Pat. Anal. & Mach. Intell.*, 11,674-693, 1989.
15. G. Olmo & L. Lo Presti, "Applications of wavelet transform for seismic activity monitoring", in *Wavelets: Theory, Algorithms and Applications*, C. Chui, L. Montefusco & L. Puccio (eds.), Academic Press, London, 1994.
16. P.J. Oonincx & S.J.L. Van Eijndhoven, "Frames, Riesz systems and MRA in Hilbert spaces", *CWI-Report PNA-R9701*, 1997.
17. O. Rioul & M. Vetterli, "Wavelets and signal processing", *IEEE Sig. Proc.*, 8, 14-38, 1991.
18. J. Samson & J. Olson, "Some comments on the descriptions of the polarization states of waves", *Geophys. J. R. astr. Soc.*, 61, 115-129, 1980.
19. R.M. Young, *An introduction to nonharmonic Fourier series*, Academic Press, New York, 1980.
20. J. Wang & T. Teng, "Identification and picking of S phase using an artificial neural network", *Bull. Seism. Soc. Am.*, 87(5), 1140-1149, 1997.