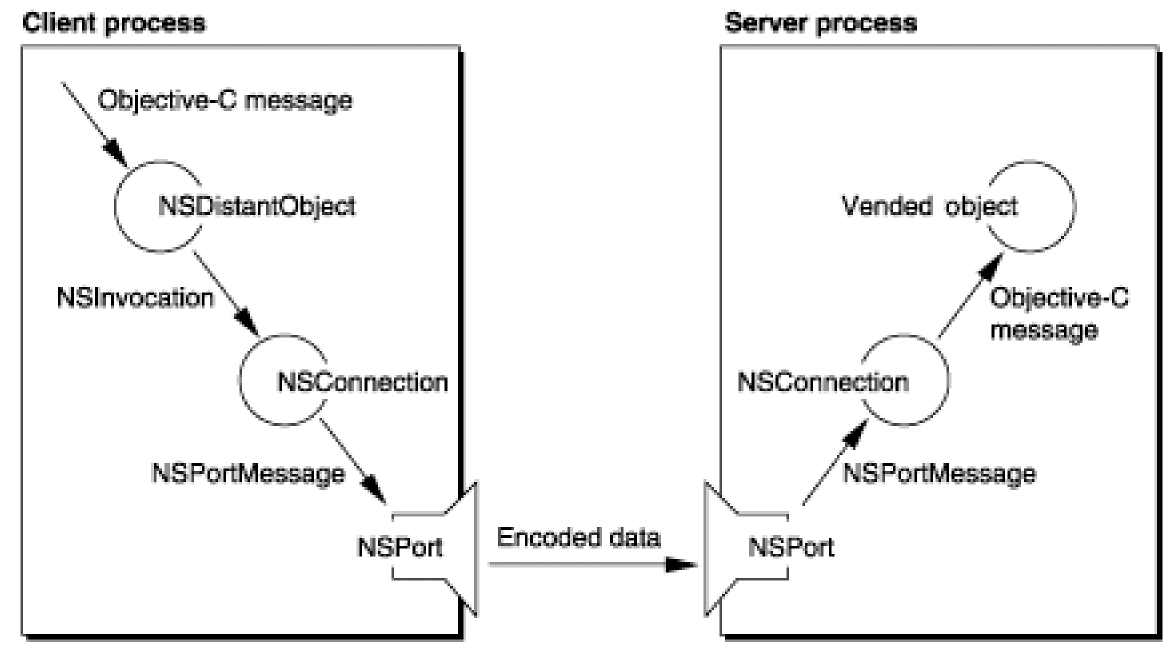
Proposal: MPI, Distributed Objects, Rendezvous, and Distributed Tasks and their use in Scientific Computing Part 1

Daniel Beatty

February 18, 2005



- All of these communication mechanisms (NSPorts, MPI, and XML-RPC) provide input and output for the distributed Turing Machine. The nodes in a distributed system (super-computer, cluster, or grid) are control units which contain a collection of states. Each object can also be considered to be a pseudo-control unit. Each one of these objects can be migrated to any of the nodes, and calls to these objects represent transition function (Turing Machine). The particulars of these transition functions, states, and control units are the subject of the Scientific Knowledge Representation and Computation Service which could very well be dissertation matter.
- NSPorts and MPI are favored to demonstrate the rudimentary transition functions and input-output operations of this distributed Turing Machine

The purpose of this document is to outline the context of the SDSS XGrid project and its specific objectives to be achieved for the Advanced Topics in Numerical Methods II. The SDSS XGrid project is part of a larger context of providing a Scientific Knowledge Base and Computational Service. For this class, there are few parts of the project which are deemed reasonable for demonstration in a Grid context. The choices of these parts include:

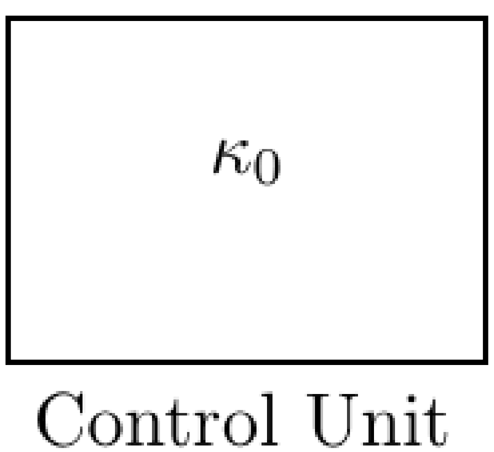
- A Wavelet transform in distributed form.
- Bonus items:
  - A template for NSPorts distributed objects with Bonjour
  - A template for distributable objects

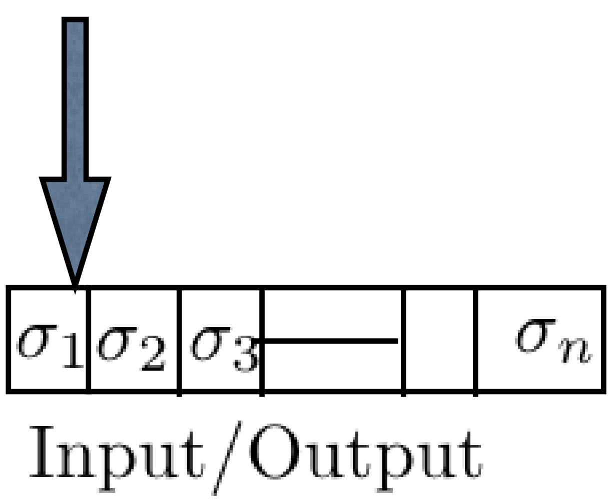
Every context of computing, including super-computing, can be described in term of the Turing Machine. The Turing Machine is a model of the simplest computing machine possible. Note that a Turing machine is a five tuple $T = \{K, \Sigma, \delta, s, h\}$ where

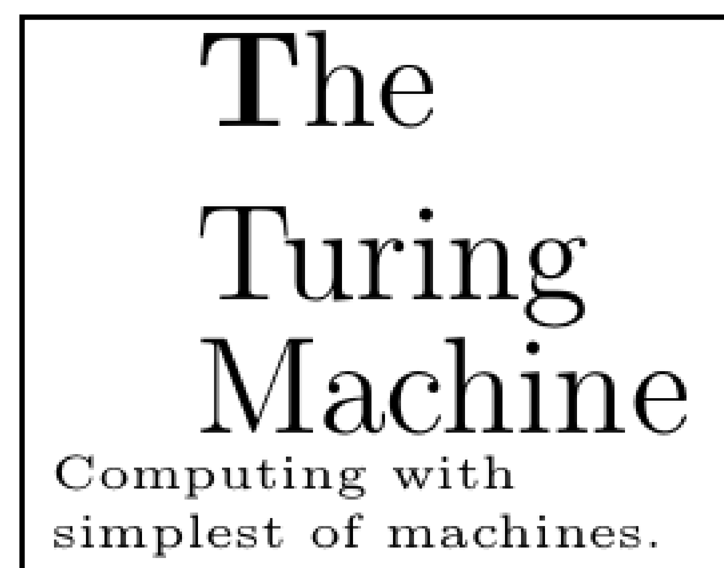
- K are the states of the Turing machine
- Σ is the alphabet of the Turing machine
- δ are transition functions on the Turing machine
- s are the starting states
- h are the halting states.

Also, a Turing machine can be viewed as a control unit with its collective of states. The transition functions tell the control unit to move from one set of states to another. A collective of states can be defined as one collective state for simplicity. In the distributed model, grid communications represents the I/O transmitting transition functions of the form

$$(K - H) \times \Sigma \rightarrow K \times (\Sigma \cup \{\leftarrow, \rightarrow\})$$







1 Using Parallel Turing Machines

There are ways use and exploit benefits in parallelizing computing. Most of the effort focus in on the input and output mechanisms connecting the control units, and the collective control of the control units. Facilities that help in generating such a collective of computing machines include:

- **Library Frameworks:** are a set of objects, protocols and modules which are contained collectively.
- **Interconnection Library (Framework)**: are library frameworks designed connect and migrate other frameworks.

In general, clusters require middle-ware such as MPI, BEEP, NSPorts, XML-RPC or other mechanism to provide the communications for their input and output. MPI, BEEP and NSPorts each form libraries which are part of a library framework. The costs of these middle-wares contribute to the engining of specific Turing machines with an emphasis on performance. During this project, BEEP was explored and discovered to be useful for the overall SDSS Knowledge Base and Computational Service. However, BEEP is not necessarily this project. MPI and NSPorts are comparable methods of distribution and are useful for this project.

There are four basic branches for all service libraries:

- The system resource manager launches publishing programs if permanent services do not already exist, or if more are possible and needed.
- The publishing program uses its associated frameworks, and publishes those objects as service objects.
- Migratory objects use marshalling/ unmarshalling methods transform themselves in reference or copy objects. These are either function or return arguments for service objects.
- Programs using service objects require discovery services and or launch services to use these services. The node from which the program itself is running may be a candidate for hosting the service libraries.

Parallel algorithms are passed a list of proxies to the other objects and protocols. One note is that parallel algorithms tend have a recursive nature to them. Thus a point of convergence is generally necessary to determine when the job stops dividing work, and performs it. Note a service that manages the launcher and coordinates the parallel algorithms would be handy.

I/O for the Parallel Turing Machine

Multiple control units, multiple inputs

In general, clusters require middle-ware such as MPI, BEEP, NSPorts, XML-RPC or other mechanism to provide the communications to exploit parallelisms. The costs of these middle-wares are subject for the performance measures section (when written). For this experiment, BEEP, NSPorts, and MPI are chosen for comparison.

Basic features that most implementations of BEEP claim the following:

- Portability
- Object Oriented Design
- Robustness and stability
- High Performance
- Multi-threaded

In essence BEEP is another message passing, and general communication protocol at the . BEEP is a layer that can be used to implement XML-RPC, NSPorts, SOAP, MPI and other service libraries. Like other web type protocols, TLS/SSL security is available. However, the XGrid implementation makes no mention of it except in its API framework.

1.1 BEEP Overview

BEEP serves as a form of glue for XGrid (zilla and zillion). It is the inter-communication framework. On OSX it is one of three frameworks for XGrid. BEEP itself stands for Block Extensible Exchange Protocol. BEEP consists of the following:

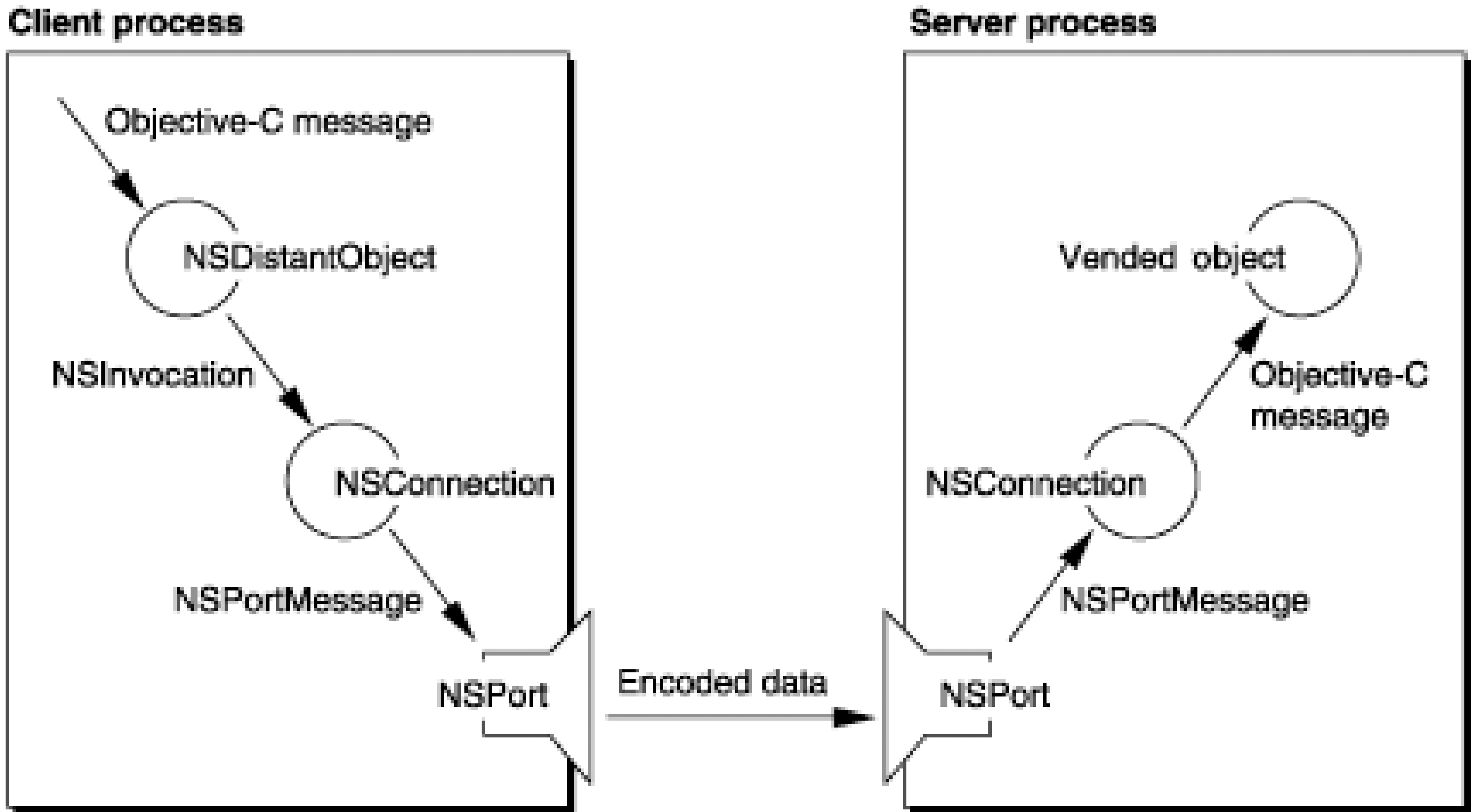
- sessions
- channels
- exchanges
- messages

[2] A session is a peer-wise, full duplex pipe. Two kinds of roles peers play are:

- Listener or Initiator
- Client / Server

BEEP connections can have multi-relations (client, server, or peer).

A channel is a full duplex pipe, and the application protocol designer specifies rules who can initiate message exchange along with a stream of type messages.



1.2 NSPorts

NSPorts is a classic method that has been available since the early days of NeXT. It is a remote object mechanism that:

- Provides objects via proxy
- Publishes services via discovery mechanisms such as Bonjour (Rendezvous)
- Provides distributed objects via marshalling and un-marshalling methods (archiving and unarchiving by Cocoa terminology).

An NSConnection object has two instances of NSPort: one receives data and the other sends data. An NSPort is a superclass to all other ports. NSMachPort uses Mach messaging and is typically used solely on the machine itself. NSSocketPorts use socket to go between machines.

- oneway void
  (client does not wait for a response.)
- in
  (A receiver is going to read the value but not change it.)
- out
  (A value is changed by the receiver by not read)
- inout
  (receiver is to both read and write the value).
- bycopy
  (argument is archived before sent and de-archived
  in the receiver's process space)
- byref
  (the argument is represented by proxy).

Introductory Wavelet Statement

- The convolution wavelet transform is a mapping from $L^2(\mathbb{R})$ defined in terms of the wavelet pair, the convolution operation on an $L^2(\mathbb{R})$ array, and an element selector.
- A wavelet pair is defined as wavelet basis function with a corresponding wavelet averaging basis function

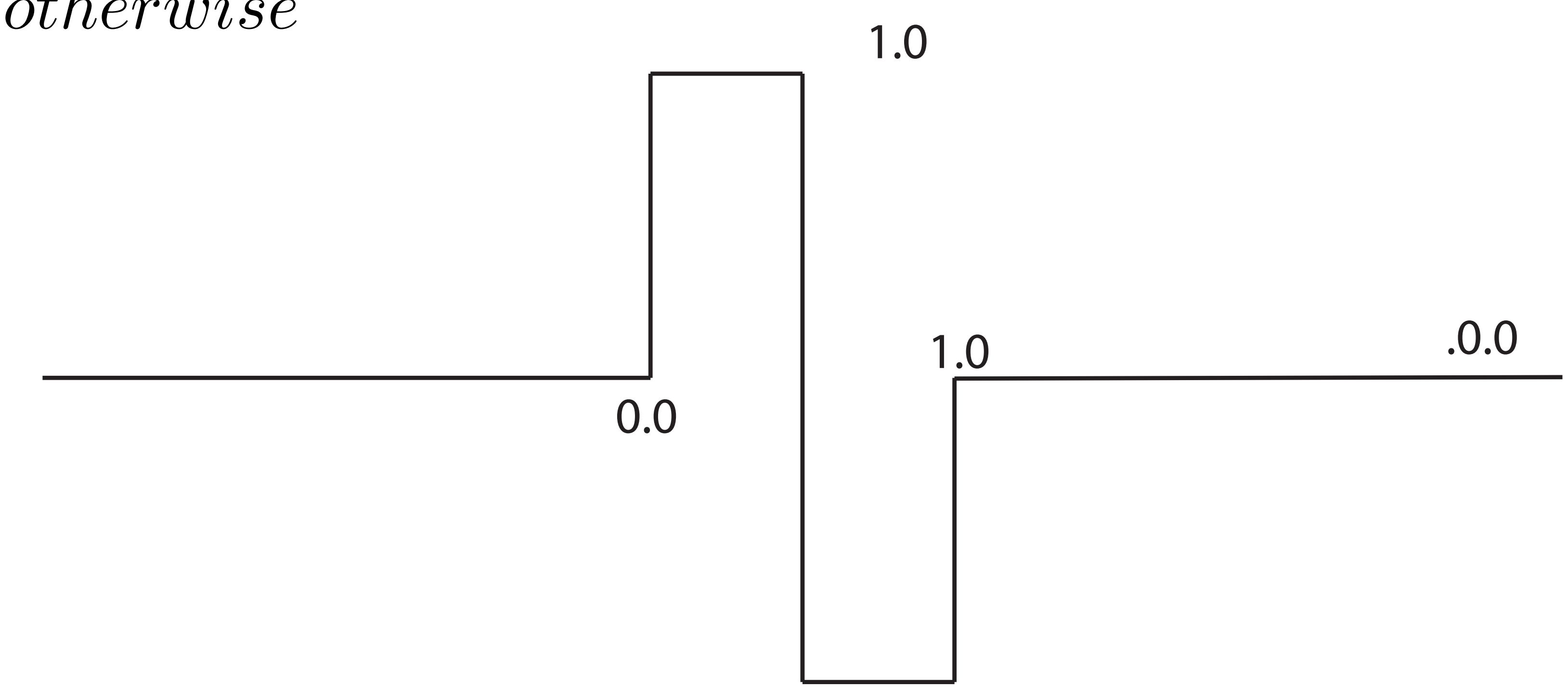
Wavelet Basis Function

The two mandatory properties of a wavelet basis function are:

- it must square integrable, and
- must have a zero average, i.e.:

$$\int_{-\infty}^{+\infty} \psi(x)\ dx = 0.$$

$$\psi(x) = \begin{cases} 1 & 0 \leq x < \frac{1}{2} \\ -1 & \frac{1}{2} \leq x < 1 \\ 0 & otherwise \end{cases} \tag{1}$$



Strict Definition of a Wavelet Basis

"A function $\psi \in L_2(R)$ is called an orthonormal wavelet if the family $\{\psi_{j,k}\}$ defined

$$\psi_{j,k}(x) = 2^{j/2}\psi(2^j x - k) \forall j, k \in Z$$

is an orthonormal basis of $L_2(R)$ where $\langle \psi_{j,k}, \psi_{l,m} \rangle = \delta_{j,l}\delta_{k,m}, \forall j, k, l, m \in Z$ and every $f \in L_2(R)$ can be written as

$$f(x) = \sum_{j,k=-\infty}^{\infty} c_{j,k}\psi_{j,k}(x)$$

where the series convergences and is

$$f(x) = \sum_{j,k=-\infty}^{\infty} c_{j,k}\psi_{j,k}(x)$$

in $L_2(R)$ such that

$$\lim_{M_1,M_2,N_1,N_2} ||f - \sum_{j=-M_2}^{N_2} \sum_{k=-M_1}^{N_1} c_{j,k}\psi_{j,k}|| = 0$$

The simplest example of orthonormal wavelets is the Haar Transform."

[10] C. Chui. *An Introduction to Wavelets.* Academic Press San Diego, 1992.

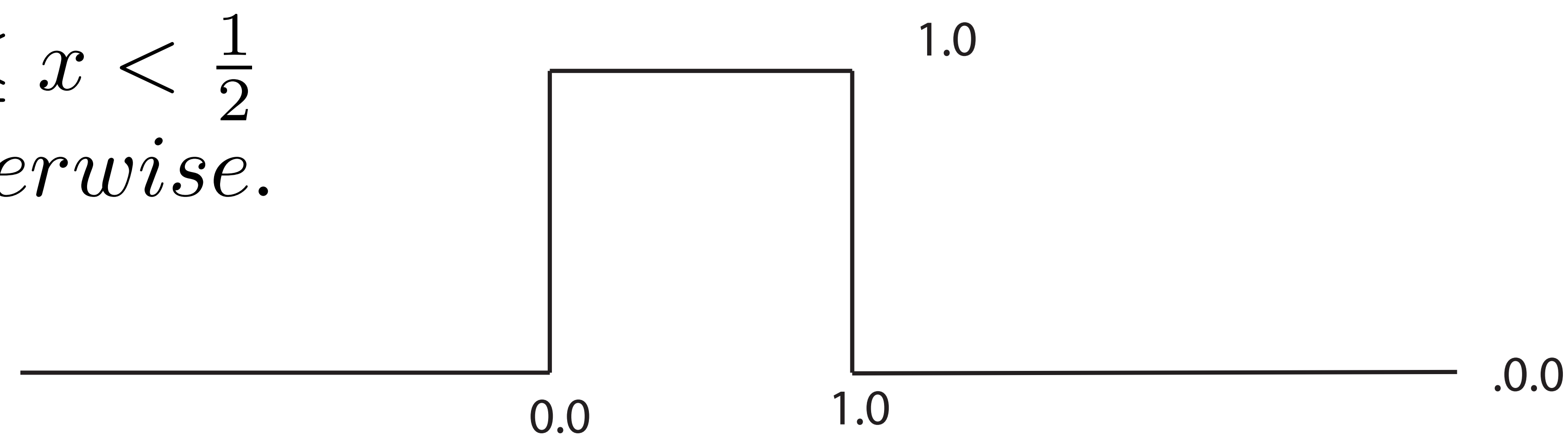
Wavelet Averaging Basis

An averaging basis function which must meet the following criteria

- be square integrable,
- satisfy the orthonormal translation-dilation property, and
- the wavelet basis function must be orthogonal with the average basis function.

Haar Averaging Basis

$$\phi(x) = \begin{cases} 1 & 0 \leq x < \frac{1}{2} \\ 0 & otherwise. \end{cases}$$



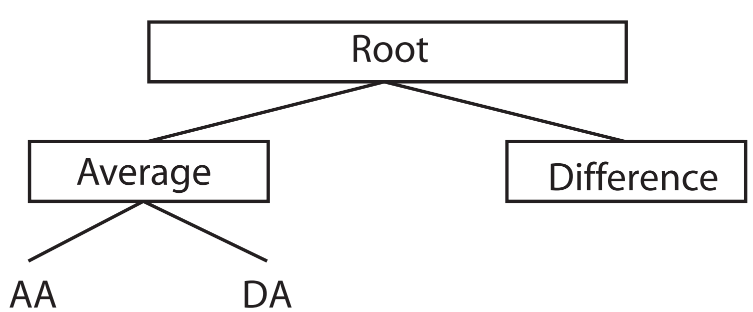
Convolution Wavelet Transform

The Convolution wavelet transform can be described using three steps

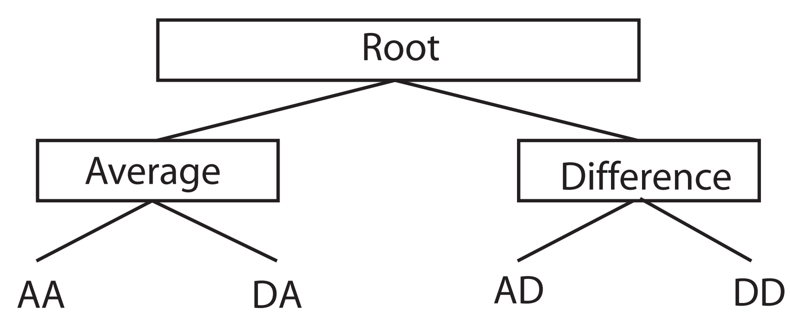
1. Convolve $A = S * \phi$ and $D = S * \psi$
2. Selectively Filter: Map $A \to A'$ and $D \to D'$
3. Concatenate: $W(S) = (A'|D')$

Multi-Resolution

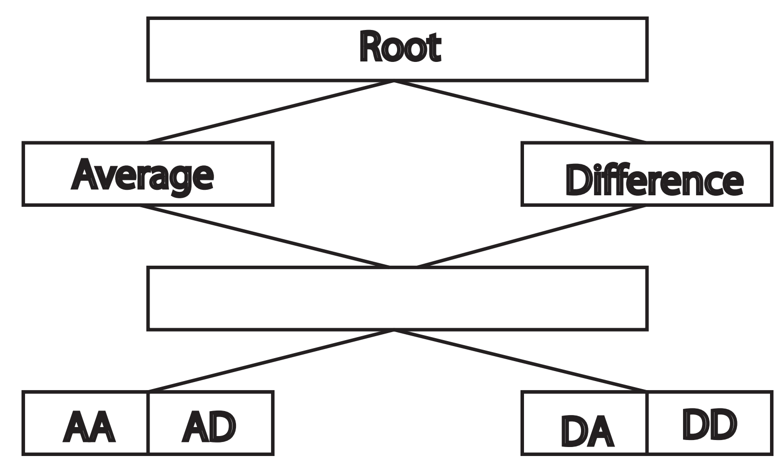
- A tree scheme is used in this thesis to represent the various forms of wavelet multi-resolution methods.
- Multi-Resolution as it applies to wavelets has three basic forms.



Multi-Resolution Analysis

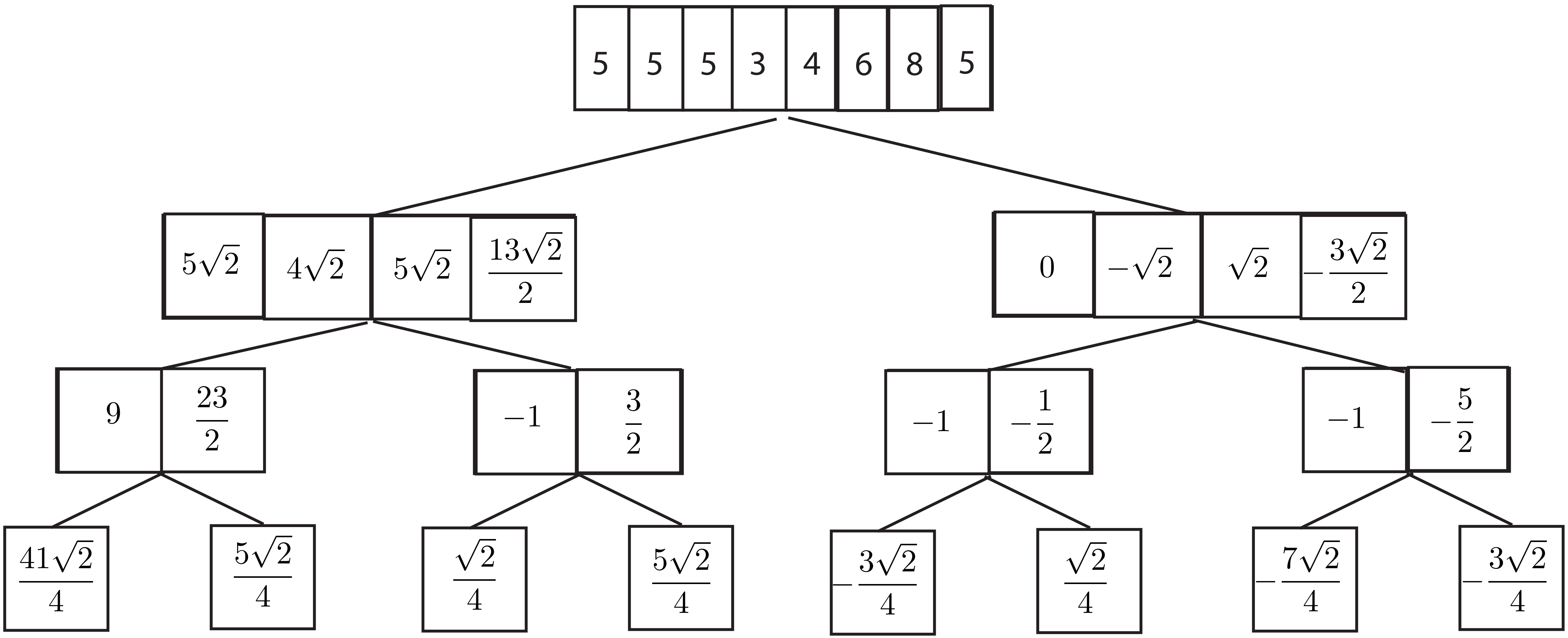


Multi-Resolution Expansion

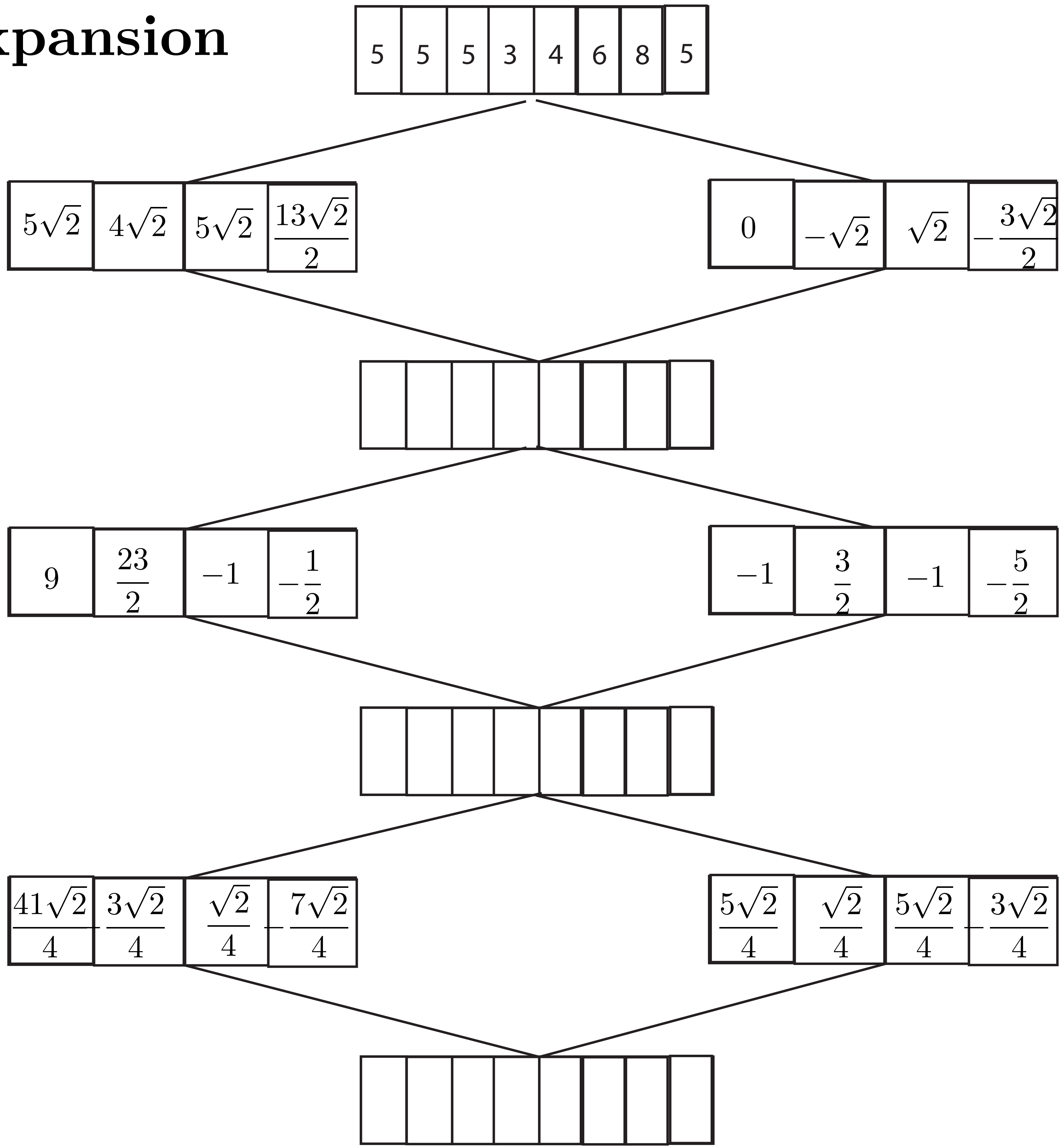


$\psi^{\mathbf{n}}$ **Expansion**

Numerical Example MRE



$\psi^{\mathbf{n}}$ Expansion



2-D Wavelet Transform

Original

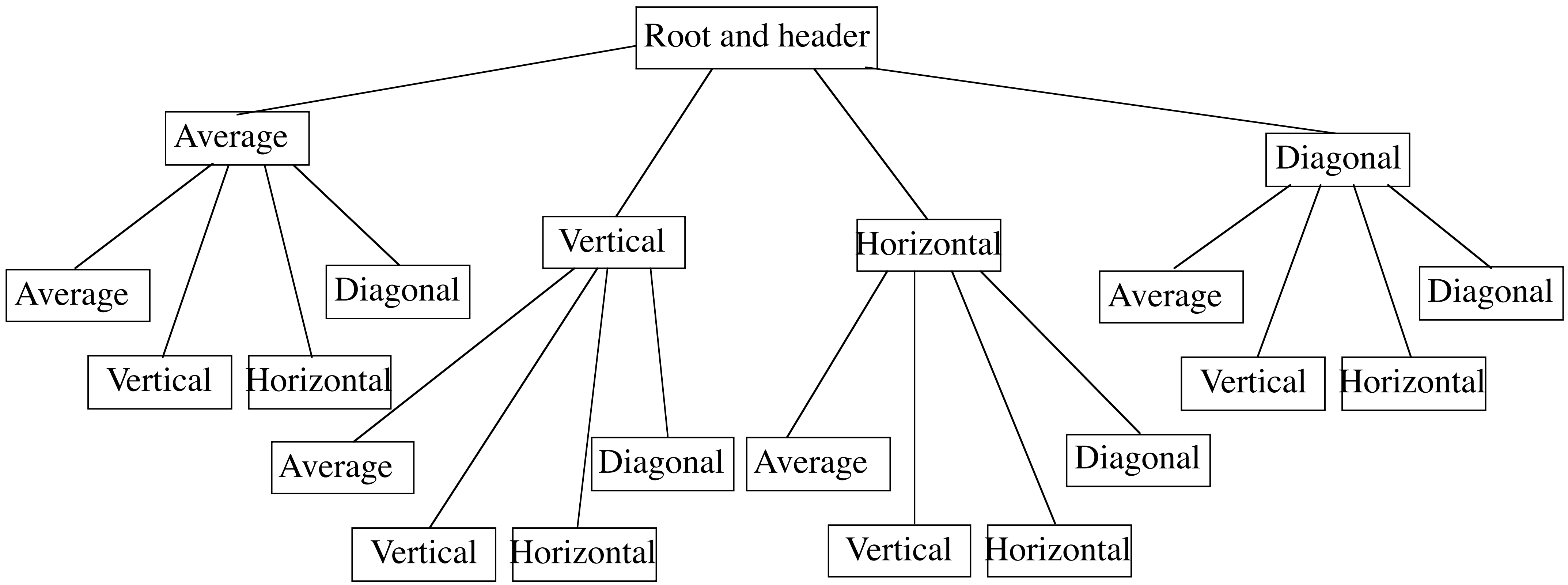
Transformed

Multi-Resolution

There exist Multi-Resolution for 2-D structures such as matrices. Those methods are:

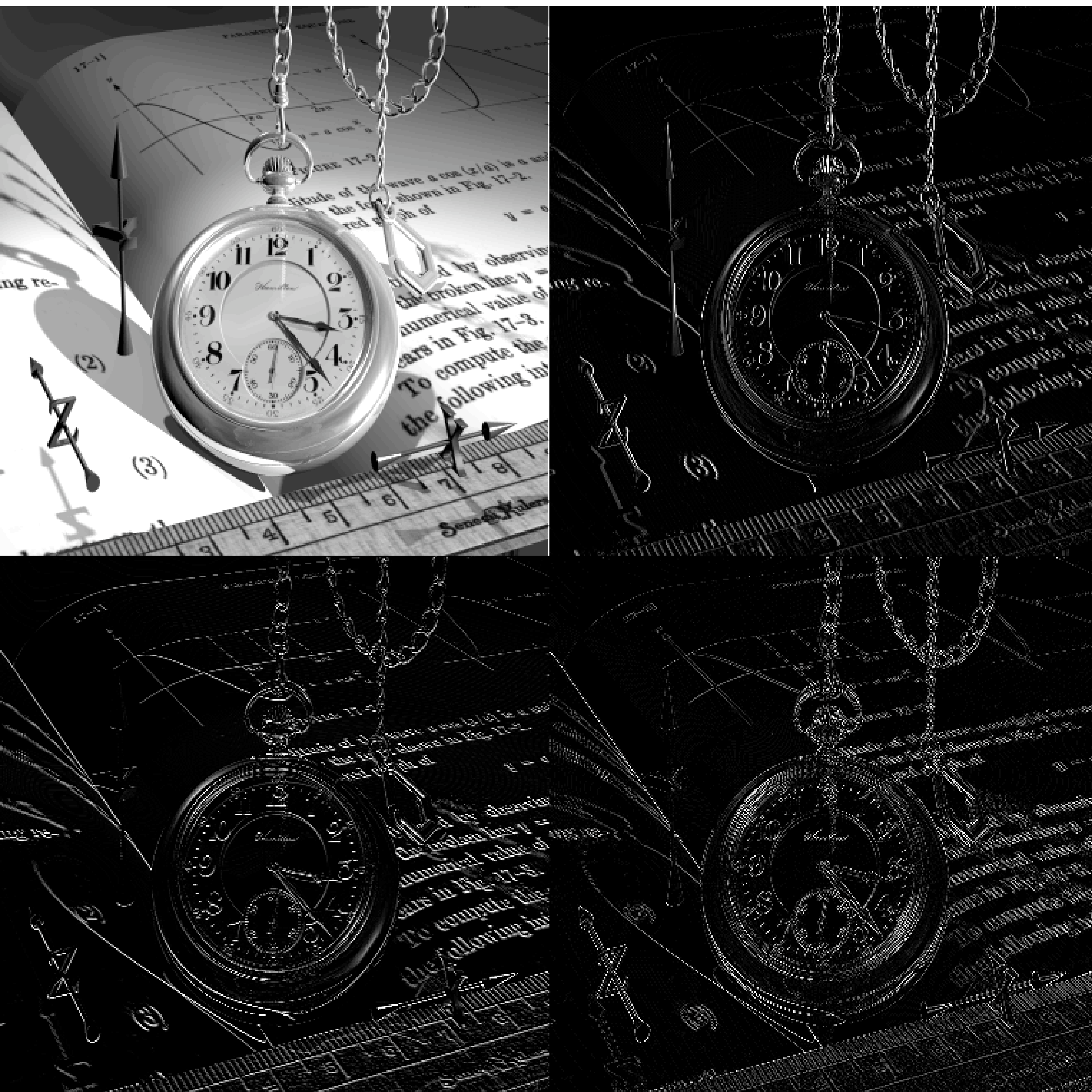
- 2- D Multi-Resolution Analysis
- 2-D Multi-Resolution Expansion
- 2-D ψ^n expansion

2-D Multi-Resoltion Expansion

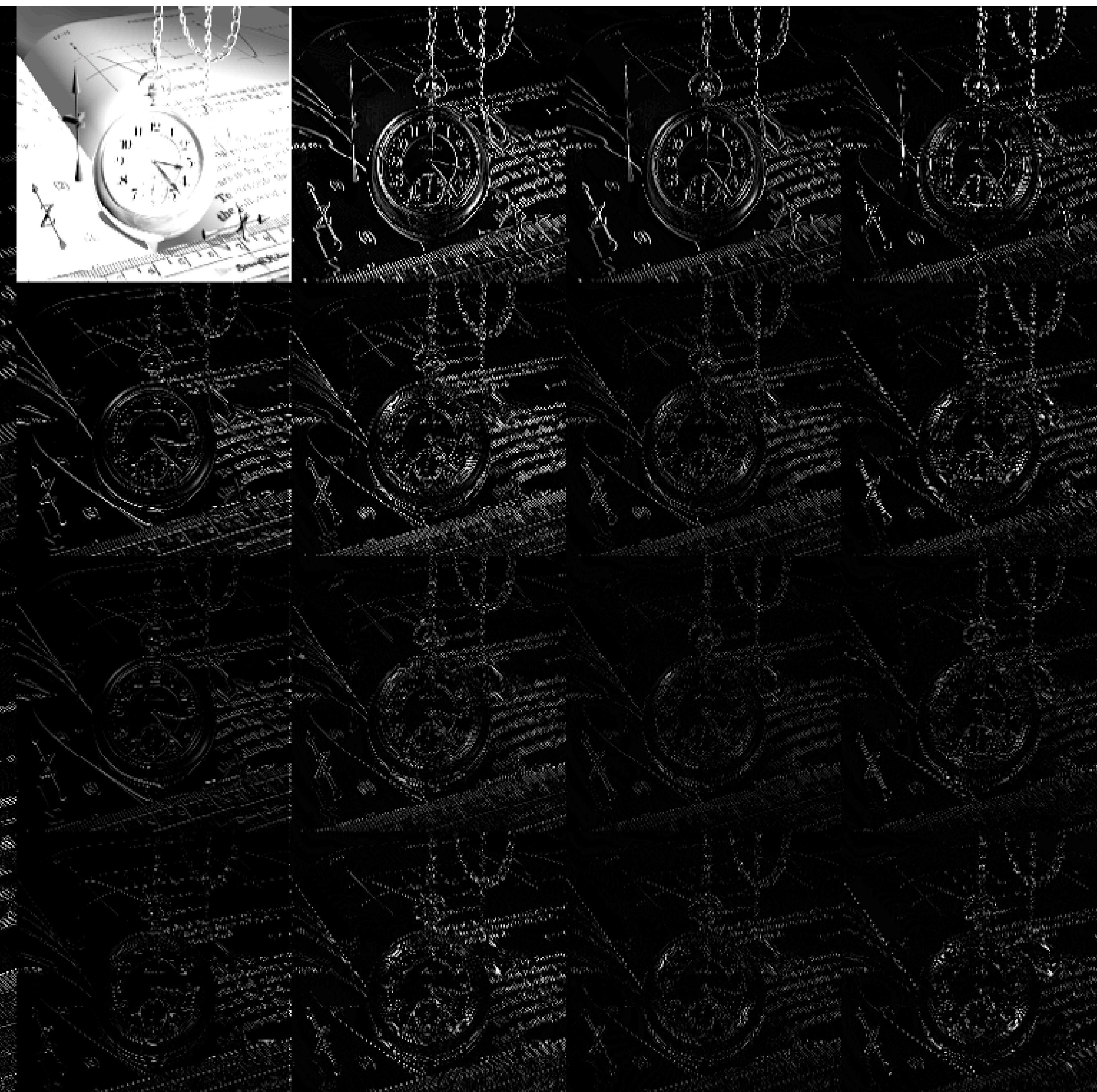


2-D MR Expansion

One Resolution



Two Resolutions



$\psi^{\mathbf{n}}$ Expansion

Algorithm 1 Wavelet Transform: MRE with queue controlled visits of the Quad Tree

Require: Wavelet Transform, MRE, and wavelet pair (ψ and ϕ)
Require: Matrix S
  Load S into temporary matrix T
  for $i = 1$ to n **do**
    $T \xrightarrow{\psi_c} X$
    $X \xrightarrow{\psi_r} T$
  end for
  return T as the transformed matrix

ψ^{n} Expansion

One Resolution

Two Resolutions

Comparison of Expansions

$\psi^{\mathbf{n}}$ **Expansion** **MR Expansion**