2016

A Wavelet Representation for a Fractional Brownian Motion

Robert Mann
University of Alabama at Birmingham

A WAVELET REPRESENTATION FOR A FRACTIONAL BROWNIAN MOTION

by

ROBERT ALFORD MANN

PAUL JUNG, COMMITTEE CHAIR
JOHN MAYER
SHANNON STARR

A THESIS

Submitted to the graduate faculty of The University of Alabama at Birmingham,
in partial fulfillment of the requirements for the degree of
Master of Science

BIRMINGHAM, ALABAMA

2016

A WAVELET REPRESENTATION FOR A FRACTIONAL BROWNIAN MOTION

ROBERT ALFORD MANN

MATHEMATICS

Abstract

Fractional Brownian Motion is a Gaussian Random Process with a covariance function that depends on a Hurst Parameter $H \in (0, 1)$. In this thesis a construction of a Standard Fractional Brownian Motion motivated by the reproducing kernel Hilbert space corresponding to the covariance function of a Standard Fractional Brownian Motion will be presented. This construction will make use of the Haar Wavelet Basis in a similar way to a classical construction of a Standard Brownian Motion. Certain facts about this construction will be proven for $H \geq \frac{1}{2}$.

Table of Contents

Abbreviations and Symbols

m and dx will stand for Lebesgue measure

$\mathbb{R}$, $\mathbb{N}$, and $\mathbb{Q}$ will stand for the real numbers, the natural numbers, and the rational numbers respectively

$\mathscr{B}(X)$ will be the Borel sigma algebra on $X \subset \mathbb{R}$

$\stackrel{d}{=}$ will mean *equal in distribution*

$:=$ will be used when an object is being defined

$L^2(\Omega) := \{f : \Omega \to \mathbb{R} : \int_\Omega |f|^2 d\mathbb{P} < \infty\}$

$\mathscr{H}$ will be used to denote a Hilbert Space

$\{\cdot\}$ will be used to denote a set of objects

1 Gaussian Processes

Random variables with Gaussian distributions, or *Gaussian Random Variables*, are very important in probability theory. Normal random variables and Gaussian random variables are the same mathematical object, so the terminology may be used interchangeably in this paper. Many of the aspects of Gaussian Random Variables are well documented and include, but are not limited to, surprising results such as The Central Limit Theorem. The definition of this important category of random variables is as follows:

Definition 1. *An n-dimensional random vector X is called Gaussian with expectation vector $m \in \mathbb{R}^n$ and a covariance matrix $\Sigma \in \mathbb{R}^{n\times n}$ if it has a density function of:*

$$\frac{1}{(2\pi)^{\frac{n}{2}}}(\det \Sigma)^{-\frac{1}{2}}e^{-\frac{1}{2}(x-m)'\Sigma^{-1}(x-m)} \textit{for all } x \in \mathbb{R}^n. \tag{1.1}$$

A Gaussian random vector of dimension one is called a Gaussian Random Variable. The characteristic function of an n-dimensional Gaussian random vector is:

$$\phi(t) = E(e^{ixt}) = e^{it'm-\frac{1}{2}(t'\Sigma t)}, t \in \mathbb{R}^n. \tag{1.2}$$

A characteristic function of a random vector is uniquely determined by its distribution and vice versa, so it is often useful to use characteristic functions to prove that a random variable is Gaussian. In the Gaussian case the characteristic function gives us one more bit of useful information. The characteristic function of a Gaussian random vector is determined by only two pieces of information, namely the mean vector m and the covariance matrix Σ. Thus it follows that a Gaussian random vector is uniquely determined by its mean vector and covariance matrix, which is an indispensable fact. The definition of a Gaussian Random Process follows naturally from the definition of a Gaussian Random Vector.

Definition 2. *A random process* $\{X_n\}_{n\in\mathbb{N}}$ *is called a Gaussian Random Process if the joint distribution of every finite collection of marginals is Gaussian.*

This paper will be primarily concerned with one specific kind of Gaussian process, called a *Fractional Brownian Motion*, but before we define what a Fractional Brownian Motion is we must first cover a few more fundamental facts about Gaussian Processes. It follows immediately from the definition that a single member of a Gaussian Process is a Gaussian Random Variable. In order to proceed we must define the mean and covariance functions for a random process, since they play an important part in Gaussian Random Processes.

Definition 3. *For any random process* $\{X_n\}_{n\in\mathbb{N}}$ *the mean function is a function* $\mu : \mathbb{N} \to \mathbb{R}$ *defined as* $\mu(n) = E(X_n)$ *and the covariance function is a function* $\sigma : \mathbb{N} \times \mathbb{N} \to \mathbb{R}$ *defined as* $\sigma(n, i) = E((X_n - E(X_n))(X_i - E(X_i)))$.

This definition can be a highly convenient definition to work with. It is obvious that if the collection of random variables are independent then $\sigma(n, i) = 0$ when $n \neq i$ and $\sigma(n, n) = Var(X_n)$. This domain of the mean and covariance function in the above definition can be generalized to any countable set of elements, which we will take advantage of below. One of the most remarkable facts about Gaussian Random Processes is that they are uniquely determined by their mean and covariance equations. This means that one only needs two pieces of information to obtain a Gaussian Process, which is an extremely useful feature of these types of random processes.

Theorem 1.1. *Given a function* $\mu : \mathbb{Q} \to \mathbb{R}$ *and a positive definite function* $\sigma : \mathbb{Q} \times \mathbb{Q} \to \mathbb{R}$*, there exists a Gaussian process* $\{X_q\}_{q\in\mathbb{Q}}$ *where* $E(X_q) = \mu(q)$ *and* $Cov(X_{q_1}, X_{q_2}) = \sigma(q_1, q_2)$ *and this Gaussian process is unique up to distribution.*

Proof. We need to define a proper sample space Ω, σ-algebra on Ω, and a probability measure on Ω. We will start by taking $\Omega = \mathbb{R}^{\mathbb{Q}}$, the function space from $\mathbb{Q}$ to $\mathbb{R}$. Let $\mathscr{F}$ be the σ-algebra defined by sets of the type $C = \{\omega_{\vec{q}} \in \mathbb{R}^n : \omega_{q_1} \in B_1 \dots \omega_{q_n} \in B_n, B_i \in \mathscr{B}(\mathbb{R})\}$. The members of this σ-algebra are typically called "cylinder subsets" of Ω. Define a collection of probability measures $\{\mathbb{P}_{\vec{q}} : \Omega \to \mathbb{R}, \vec{q} \in \mathbb{Q}^n\}$ by

$$\mathbb{P}_{\vec{q}}(C) = \int_{B_1,\dots,B_n} \frac{1}{(2\pi)^{\frac{1}{2}}} (\det \Sigma)^{-\frac{1}{2}} e^{-\frac{1}{2}(x-m)'\Sigma^{-1}(x-m)} dx \tag{1.3}$$

where Σ is an $n \times n$ dimensional matrix with components defined by $(\Sigma)_{i,j} = \sigma(q_i, q_j)$ and m is an n-dimensional vector with components defined by $(m)_i = \mu(q_i)$.

It can be easily shown that the collection of probability measures $\{\mathbb{P}_{\vec{q}}\}$ follows the consistency conditions needed to invoke the Kolmogorov extension theorem. By Kolmogorov's extension theorem there exists a unique infinite dimensional extension of $\{\mathbb{P}_{\vec{q}}\}$ on the probability space $(\Omega, \mathscr{F})$ for which $\{\mathbb{P}_{\vec{q}}\}$ are the marginals. Now define a stochastic process $\{X_q\}_{q \in \mathbb{Q}}$ by setting $X_{q_i} = \omega_{q_i}$. The joint distribution of every finite number of these random variables is given above by (1.3). Thus this stochastic process is, by definition, Gaussian. As mentioned above, since multivariate Gaussian random variables are uniquely determined by their Mean and Covariance matrix, this insures that the distribution of every linear combination of X_{q_i}'s is unique, and thus the process is unique. $\square$

This theorem says that whenever we define a Gaussian process we only need a mean and covariance function in order to properly do it. For the rest of this paper we will be working with Gaussian processes, so this theorem will be used extensively. One famous example of a Gaussian process is a Brownian Motion.

Definition 4. *Consider the campact interval* $[0, T]$. *By Theorem 1.1 construct a Gaussian process* $\{B_{t_q}\}_{t_q \in \mathbb{Q} \cap [0,T]}$ *with*

1. $E(B_{t_q}) = 0$

2. $E(B_{t_q}B_{s_q}) = min\{t_q, s_q\}$

3. $B_0 = 0$.

Extend this random process to $[0,T]$ *such that* $\{B_t\}_{t\in[0,T]}$ *is continuous in* $[0,T]$ *with probability one.* $\{B_t\}_{t\in[0,T]}$ *is called a Standard Brownian Motion.*

It is not immediately obvious whether or not a continuous extension exists, but this result is well known[Ste01]. Any Brownian Motion with $B_0 = 0$ can be represented as $\mu t + \sigma B_t$ where B_t is a standard Brownian Motion, which is also called a Brownian Motion *With Drift*(starting at the origin). Brownian Motion is one of the most extensively studied Gaussian process which manifests itself in a surprising number of applications. It is fortunate that Brownian Motion is a useful process, because it also is the easiest Gaussian process to perform mathematical analysis on. Apparently, every continuous Gaussian process with stationary and independent increments must be a Brownian Motion. The useful properties of Brownian Motion are well documented and can fill entire textbooks, but for the remainder of this paper we will be primarily concerned with a generalization of Brownian Motion called *Fractional Brownian Motion.*

2 Fractional Brownian Motion

In applied mathematics one of the most useful properties of Brownian Motion is its independent increments. However, as is frequently the case, this convenient property is often the most questionable when Brownian Motion is used in applied models. Brownian Motion does have useful properties other than independent increments and it has been used in applied fields to such great success that it should also not be disregarded simply for this one flaw. Ideally one would desire to create a class of random processes related to each other in such a manner that they generalize Brownian Motion in the sense that Brownian Motion is included in this class of random processes. One such class of random processes that fits these requirements is the class of Fractional Brownian Motions.

Definition 5. *Consider the compact interval* $[0,T]$*. Given* $H \in (0,1)$ *construct a Gaussian process* $\{B_{t_q}^H\}_{t_q \in \mathbb{Q} \cap [0,T]}$ *with*

1. $E(B_{t_q}^H) = 0$

2. $E(B_{s_q}^H B_{t_q}^H) = \frac{1}{2}(t_q^{2H} + s_q^{2H} - |t_q - s_q|^{2H})$

3. $B_0^H = 0$.

Extend this to a Gaussian process $\{B_t^H\}_{t \in [0,T]}$ *such that* B_t^H *is continuous in* $[0,T]$ *with probability one.* $\{B_t^H\}_{t \in [0,T]}$ *is a Standard Fractional Brownian Motion.*

As with Brownian Motion, it is not immediately obvious that such a continuous extension exists, but again this is a well known result[Bia08]. In this case H is called the Hurst Parameter. Just as with Brownian Motion, general Fractional Brownian Motions starting at the origin can be represented as shifted and augmented Standard Fractional Brownian Motions. It immediately follows from the definition that $E(B_s^{\frac{1}{2}} B_t^{\frac{1}{2}}) = \frac{1}{2}(t + s - |t - s|) = \min\{s,t\}$. Therefore a Fractional Brownian Motion with Hurst Parameter $H = \frac{1}{2}$ is actually a Brownian Motion, which is one of desired properties of this class of random processes. The question of existence will be addressed through a specific construction during the Main Results Section of this thesis. The next theorem helps to establish two very important properties of Fractional Brownian Motion.

Theorem 2.1. *If* $\{B_t^H\}_{t \in [0,T]}$ *is a Standard Fractional Brownian Motion with Hurst Parameter* H *then*

1. $\{B_t^H\} \stackrel{d}{=} \{a^{-H} B_{at}^H\}$ *(Self-Similarity)*

2. $\{B_t^H\}$ *has stationary increments.*

Proof. We prove 1 first. Obviously $\{a^{-H} B_{at}^H\}$ is a Gaussian process, so by Theorem 1.1 we only need to prove that the first and second moments of the two processes are the same. It is immediate that, for a fixed t, $E(B^H_t) = 0 = a^{-H}E(B^H_{at}) = E(a^{-H}B^H_{at})$, so the only remaining thing to check is the covariance equation. Indeed, for $s \neq t$

$$\begin{aligned} E(B^H_t B^H_s) &= \frac{1}{2}(t^{2H} + s^{2H} - |t-s|^{2H}) \\ &= \frac{a^{-2H}}{2}((at)^{2H} + (as)^{2H} - |at-as|^{2H}) \\ &= E(a^{-2H}B^H_{at}B^H_{as}) \end{aligned}$$

which proves the first statement.

The second statement can be proved similarly. Consider arbitrary finite sets $\{B_{t_{i+1}} - B_{t_i}\}^n_{i=1}$ and $\{B_{t_{i+1}+s} - B_{t_i+s}\}^n_{i=1}$. Since Fractional Brownian Motion is a Gaussian Process the joint distribution of these random variables are Gaussian, so in order to prove stationarity we only need to prove their first and second moments are the same. Both of these sets obviously have zero mean, so we only need to prove that the covariance matrices are the same. Assuming $t_n \geq v > u \geq t > r \geq t_1$:

$$\begin{aligned} &E((B_v - B_u)(B_t - B_r)) = (v-t)^{2H} - (u-t)^{2H} + (u-r)^{2H} - (r-v)^{2H} \\ &= (v+k-(t+k))^{2H} - (u+k-(t+k))^{2H} + (u+k-(r+k))^{2H} - (r+k-(v+k))^{2H} \\ &= E((B_{v+s} - B_{u+s})(B_{t+s} - B_{r+s})). \end{aligned}$$

This proves that the covariance matrices of the above random variables are the same, and thus the increments of Fractional Brownian Motion are stationary. □

This theorem shows that, despite losing independent increments, moving from Brownian Motion to Fractional Brownian Motion maintains stationary increments and self similarity. Thus, despite losing some mathematical convenience, Fractional Brownian Motion still has some nice structure. Fractional Brownian Motion's pragmatic use and mathematical structure have caused it to become a very exciting random process in recent years.

As established above, the important aspect that makes Fractional

Brownian Motion more pragmatic than a Brownian Motion is that it allows for some dependency between its increments. The probablistic interpretation of this aspect of Fractional Brownian Motion is obvious, namely that the results of a Fractional Brownian Motion during one increment of time will affect how it moves in future increments of time. This intepretation seems to correspond to many phenomena we see in many academic subjects, including economics, finance, and a wide range of physical applications.

Depending on the Hurst Parameter being used, the intervals of Fractional Brownian Motion can have different kinds of dependency structures. If $H > \frac{1}{2}$ then the Fractional Brownian Motion exhibits what is called *long range dependence*, which means that the autocovariance function between increments decays at a slower-than-exponential rate. For $H < \frac{1}{2}$ *short range dependence* is exhibited. One can actually perform statistics to figure out, within an error term, the Hurst parameter of a time series in order to use the most predictive model. In finance, for example, it has been demonstrated that in many financial markets the movement across intervals of time can actually exhibit long range dependence[EHK03].

The main concern of this thesis will be constructing a specific representation of a Fractional Brownian Motion with $H > \frac{1}{2}$. This will be achieved by using a specific kernel function(in the sense of a Reproducing Kernel Hilbert Space) to achieve the desired covariance function for the Fractional Brownian Motion. We will also be using the Haar Wavelet Basis in [0,1] to construct the Fractional Brownian Motion, which will be explored in the next section.

3 Wavelets

Very often doing functional analysis begins with finding an orthonormal basis, which always exists in a separable Hilbert Space. In this paper we will be concerned with orthonormal bases in $L^2[0,T]$. The classical basis for $L^2[0,T]$ is the Fourier basis, $f_n(x) = e^{inx}$. The Fourier basis is in L^2 and forms a complete set in the space of finite second moment Lebesgue integrable functions with domain $[0, T]$. The Fourier basis is, for both mathematical and historical reasons, the first orthonormal basis most students learn.

The Fourier basis is undoubtedly a useful basis, but it also poses a few problems once one tries to apply the mathematics to real world subjects. Orthonormal bases in a Hilbert Space are useful because they allow one to approximate any function in the Hilbert Space with a linear combination of the basis elements. In the case of the Fourier basis every function $f \in L^2[0, 1]$ can be represented as:

$$f(x) = \sum_{n=1}^{\infty} \left(\int_0^1 f(t) e^{-in2\pi t} dt \right) e^{in2\pi x}.$$

The terms in the parentheses in the above equation are called the Fourier coefficients and the above equation implies that all one needs in order to completely define $f \in L^2$ is to calculate its Fourier coefficients. The Fourier basis has been extensively used for signal processing, due to the sinusoidal nature of its coefficients, and it seems to be well suited to this task. However, the Fourier coefficients do not exhibit what is called time-localization.

In the language of signal processing, we typically think of a function meant to model a signal as both a function of frequency and time. The above Fourier coefficients can be intepreted as an attempt to integrate out the time variable from the signal function so that one can identify the function using only the frequency space. What if we want to know how the signal acts as a function of frequency, but only over a specific time interval? This is where the classic Fourier coefficients lose a bit of explanatory power since they are not defined over specific time intervals but rather over the entire time-domain.

As mentioned above, this problem is typically coined as a time-localization problem. There are easy ways to fix this, and one can even fix it without leaving the language of Fourier Analysis. Classically this problem has been fixed by introducing what are called window functions, which are shifts of a compactly supported function that are designed to localize the Fourier coefficients at predetermined time intervals. However, this approach can be clunky since the window functions often must be chosen based on the scenario one is faced with.

One would hope that there is an orthonormal basis in $L^2[0,T]$ that has the time localization property built into it by design. Wavelet theory is a theory that is designed partially to address this issue. Wavelet theory is a very broad theory which will not be explored in this thesis, but if one wants a comprehensive analysis then one can look in [ID92]. In this paper we will be using a specific wavelet basis called the Haar wavelet basis.

Definition 6. *Define*

$$P(x) := \begin{cases} 1 & \textit{if } 0 \leq x \leq \dfrac{1}{2} \\ -1 & \textit{if } \dfrac{1}{2} < x \leq 1 \\ 0 & \textit{elsewhere} \end{cases}$$

as the Haar Mother Wavelet. *For each $n \in \mathbb{N}$ associate j, k such that $n = 2^j + k$. The* Haar Basis *is a sequence $\{P_n\}$ defined as:*

$$\begin{aligned} P_0(x) &:= \begin{cases} 1 & \text{if } 0 \leq x \leq 1 \\ 0 & \text{elsewhere} \end{cases} \\ P_n(x) &:= 2^{\frac{j}{2}} P(2^j x - k) \\ &= \begin{cases} 2^{\frac{j}{2}} & \text{if } \dfrac{k}{2^j} \leq x \leq \dfrac{1+2k}{2^{j+1}} \\ -2^{\frac{j}{2}} & \text{if } \dfrac{1+2k}{2^{j+1}} < x \leq \dfrac{1+k}{2^j} \\ 0 & \text{elsewhere.} \end{cases} \end{aligned} \tag{3.4}$$

This definition is only defined on the domain space $[0,1]$, but it can be easily extended to any closed bounded interval $[0, T]$ by simply renormalizing the mother wavelet. If we assume, for the moment, that the Haar wavelet basis is in fact a basis in $L^2[0, T]$ then it can be easily seen that the wavelets are only defined on local portions of $[0, T]$. Thus if we use them in a similar manner to the way the Fourier basis is used in the Fourier coefficients then we have the desired time-localization property. In order to complete this analysis we must have the following theorem, the proof of which can be found in [ID92]:

Theorem 3.1. *The Haar basis forms a complete orthonormal set in $L^2[0, 1]$.*

We will be using the Haar Wavelets in our next section as a complete orthonormal basis in order to construct a version of a Fractional Brownian Motion. Proving that the proposed random process is in fact a Fractional Brownian Motion will require invoking Parseval's identity, so the above theorem is of the upmost importance to our construction.

4 Reproducing Kernel Hilbert Spaces

In this section we will introduce the basic theory of Reproducing Kernel Hilbert Spaces. We will not present proofs for some of the preliminary theorems in this section, they will be used to provide a solid theoretical basis for later in the section. It turns out that Reproducing Kernel Hilbert Spaces have close relationships to Gaussian Processes, as we will explore later in this section.

Definition 7. *Given a set $[0, T] \subset \mathbb{R}$ and a symmetric, positive definite $R : [0, T] \times [0, T] \to \mathbb{R}$ the unique Hilbert Space $\mathscr{H}(R)$ is called a Reproducing Kernel Hilbert Space with Reproducing Kernel R if*

1. $R(\cdot, t) \in \mathscr{H}(R)$
2. $\langle g, R(\cdot, t)\rangle_{\mathscr{H}(R)} = g(t)$ *for every* $g \in \mathscr{H}(R)$

It turns out that every $g \in \mathscr{H}(R)$ can be represented as $g(x) = \displaystyle\sum_{i=1}^{N} a_i R(x, t_i)$ or the limits of these functions under the norm

$||g||^2 = \sum_{i,j=1}^{N} a_i R(t_j, t_i) a_j$. Note that each member of the RKHS is defined pointwise. The RKHS for a given symmetric, positive definite function always exists, is unique defined on $[0, T]$, and it also has a close relationship to a Gaussian process with covariance function R, which is seen in the next theorem.

Theorem 4.1. *If $\{X_t\}_{t\in[0,T]}$ is a real valued Gaussian Process with covariance function R_X and mean function $m \in \mathscr{H}(R_X)$ then there exists a map $\Phi : \mathscr{H}(R_X) \to L^2(\Omega_X, \mathbb{P})$ which satisfies the following properties for all $t \in [0, T]$ and $g, h \in \mathscr{H}(R_X)$:*

1. *Φ is linear, one to one, and onto,*
2. *$\Phi(R_X(\cdot, t)) = X(t)$, and*
3. *$E\,(\Phi(g)) = \langle m, g\rangle_{\mathscr{H}(R_X)}$, $cov\,(\Phi(g), \Phi(h))= \langle h, g\rangle_{\mathscr{H}(R_X)}$.*

It follows pretty easily from this theorem that $cov\,(X(t), \Phi(g)) = cov\,(\Phi(R_X(\cdot, t)), \Phi(g)) = \langle R_X(\cdot, t), g\rangle = g(t)$, so we can immediately characterize every function in the RKHS by the mapping Φ. For the purpose of this paper we will be concerned with the RKHS associated with the covariance function of a Standard Fractional Brownian Motion $\mathscr{H}(R_H)$ for a fixed H, but we do need two more elementary results from RKHS theory before we sharpen our focus back to Fractional Brownian Motion.

Theorem 4.2. *If $X = \{X(t)\}$ is a centered Gaussian Process with covariance function R_X and there exists functions $\{f_t, t \in [0, T]\}$ such that $R_X(s, t) = \int_{[0,T]} f_s(x) f_t(x) dx$ then*

$$X(t) = \int_{[0,T]} f_t(x) dB_x \tag{4.5}$$

where B_x is a Standard Brownian Motion on $[0, T]$.

These fundamental theorems about RKHS's have more theoretical forms but since we are concerned with $\mathscr{H}(R_H)$ and the probability space

$([0,T],\mathscr{B}([0,T]),m)$ the theorems have been tailored for that purpose. By these previous theorems we want to find a sequence of functions $\{f_t, t\in[0,T]\}$ such that $\frac{1}{2}(t^{2H}+s^{2H}-|t-s|^{2H}) = \int_{[0,T]} f_s f_t dx$. In the next lemma we will present a possible candidate for these f_t's and prove that they satisfy the above requirements. Before we prove this lemma we need to recall that the *Beta Function* is defined as:

$$\beta(y,z) := \int_0^1 x^{y-1}(1-x)^{z-1}dx.$$

Lemma 4.3. *Define the two expressions for $T\geq t>s>0$:*

$$K_H(t,s) := c_H s^{\frac{1}{2}-H}\int_s^t (u-s)^{H-\frac{3}{2}}u^{H-\frac{1}{2}}du \tag{4.6}$$

$$c_H := \left(\frac{H(2H-1)}{\beta(2-2H,H-\frac{1}{2})}\right)^{\frac{1}{2}}. \tag{4.7}$$

Then, for $H>\frac{1}{2}$

$$\frac{1}{2}(t^{2H}+s^{2H}-(t-s)^{2H}) = \int_0^T 1_{(0,s)}(v)K_H(s,v)1_{(0,t)}(v)K_H(t,v)dv.$$

Proof.

We will start in our proof with this basic equality:

$$\frac{1}{2}(t^{2H}+s^{2H}-(t-s)^{2H}) = H(2H-1)\int_0^s\int_0^t |r-u|^{2H-2}dudr$$

Now we will work with $|r-u|^{2H-2}$, assume, without loss of generality, $r>u$:

$$
\begin{aligned}
(r-u)^{2H-2} &= \frac{(ru)^{H-\frac{1}{2}}}{\beta(2-2H,H-\frac{1}{2})}\left((ru)^{\frac{1}{2}-H}\beta(2-2H,H-\frac{1}{2})(r-u)^{2H-2}\right)\\
&= \frac{(ru)^{H-\frac{1}{2}}}{\beta(2-2H,H-\frac{1}{2})}\left((ru)^{\frac{1}{2}-H}(r-u)^{2H-2}\int_0^1 (1-x)^{1-2H}x^{H-\frac{3}{2}}dx\right)\\
&= \frac{(ru)^{H-\frac{1}{2}}}{\beta(2-2H,H-\frac{1}{2})}\left((r-u)^{2H-2}\int_{\frac{r}{u}}^{\infty} (zu-r)^{1-2H}z^{H-\frac{3}{2}}dz\right)\\
&= \frac{(ru)^{H-\frac{1}{2}}}{\beta(2-2H,H-\frac{1}{2})}\left(\int_0^u v^{1-2H}(r-v)^{H-\frac{3}{2}}(u-v)^{H-\frac{3}{2}}dv\right).
\end{aligned}
$$

Where in line 3 we did a substitution $z = \dfrac{r}{ux}$ and in line 4 we did the substition $v = \dfrac{r-zu}{1-z}$. So, based on the above equation, we have

$$
\begin{aligned}
&\frac{1}{2}(t^{2H}+s^{2H}-(t-s)^{2H})\\
&= \frac{H(2H-1)}{\beta(2-2H,H-\frac{1}{2})}\left(\int_0^s\int_0^t (ru)^{H-\frac{1}{2}}\left(\int_0^{r\wedge u} v^{1-2H}(r-v)^{H-\frac{3}{2}}(u-v)^{H-\frac{3}{2}}dv\right)drdu\right)\\
&= c_H^2\int_0^s v^{1-2H}\int_v^s u^{H-\frac{1}{2}}(u-v)^{H-\frac{3}{2}}du\int_v^t r^{H-\frac{1}{2}}(r-v)^{H-\frac{3}{2}}drdv\\
&= \int_0^s K_H(s,v)K_H(t,v)dv\\
&= \int_0^T 1_{(0,s)}(v)K_H(s,v)1_{(0,t)}(v)K_H(t,v)dv.
\end{aligned}
$$

□

The above lemma and Theorem 4.2 establish that

$$
B_H(t) = \int_{[0,T]} 1_{(0,t)}K_H(t,x)dB_x. \tag{4.8}
$$

Since Brownian Motion is an extensively studied process and K_H is a well defined function we now have a candidate for a new representation of a Fractional Brownian Motion. Using the Haar Wavelets introduced in Section 3 it is well known that a Standard Brownian Motion has the representation

$$B(t) = \sum_{n=0}^{\infty} Z_n \int_{[0,T]} 1_{(0,t)} P_n(x) dx$$

where $\{Z_n\}_{n \in \mathbb{N}}$ is a sequence of independent identically distributed standard normal random variables. If we take an intuitive approach then we can informally guess from this well known representation that $dB_x \approx \sum_{n=0}^{\infty} Z_n P_n(x) dx$ and thus plugging this into (4.8) we get the equation

$$B_H(t) \approx \sum_{n=0}^{\infty} Z_n \int_{[0,T]} 1_{(0,t)} K_H(t,x) P_n(x) dx.$$

Thus, while plenty remains to be proven, we know have a solid candidate for a wavelet construction of a Fractional Brownian Motion. In the next section we will prove a few facts about this random variable.

5 A Construction of a Fractional Brownian Motion

In this section we will present a construction of Fractional Brownian Motion by using the K_H functions introduced in Section 4, wavelets, and standard normal random variables. Before we introduce the construction we must prove a fact about normal random variables that we will need to use later. The following lemma and proof was inspired by a similar proof in [Ste01].

Lemma 5.1. *If $\{Z_n : n \in \mathbb{N}, n \geq 0\}$ is a sequence of independent standard normal random variables then there exists a random variable $X(\omega)$ such that*

$$|Z_n| \leq X(\omega)\sqrt{\log(n+2)}$$
$$\mathbb{P}(X(\omega) < \infty) = 1.$$

Proof. For $x \geq 1$,

$$\begin{aligned}\mathbb{P}(|Z_n| \geq x) &= \frac{2}{\sqrt{2\pi}} \int_x^\infty e^{\frac{-u^2}{2}} du \\ &\leq \frac{2}{\sqrt{2\pi}} \int_x^\infty u e^{\frac{-u^2}{2}} du \\ &= e^{\frac{-x}{2}} \frac{2}{\sqrt{2\pi.}}\end{aligned}$$

If we choose $x = \sqrt{2\alpha \log(n+2)}$ where $\alpha > 2$ and $n \geq 0$ then

$$\begin{aligned}\mathbb{P}\left(|Z_n| \geq \sqrt{2\alpha \log(n+2)}\right) &\leq \exp\left(-\alpha \log(n+2)\right) \\ &= (\sqrt{n+2})^{-\alpha}.\end{aligned}$$

By the Borel Cantelli Lemma this implies

$$\mathbb{P}(|Z_n| \geq \sqrt{2\alpha \log(n+2)} \text{ infinitely often}) = 0.$$

This implies that the random variable $X(\omega) = \sup_{n \geq 0} \dfrac{|Z_n|}{\sqrt{\log(n+2)}}$ is finite with probability 1 and $|Z_n| \leq X(\omega)\sqrt{\log(n+2)}$.

□

Recalling the definition of the Haar wavelet basis we gave in Section 3 we can state the next useful lemma:

Lemma 5.2. *Let $H_1 = L^2(\Omega, \mathcal{F}, \mathbb{P})$ be a mean zero Gaussian space. Assume $\{Z_n : n \geq 0\}$ is an orthonormal basis of standard normal random variables. $\{Z_n(\omega)P_k(x)\}_{n,k\in\mathbb{N}}$ forms an orthonormal basis in the space $L^2(\Omega \times [0,T])$.*

This is a fairly elementary result that can be found in textbooks on Hilbert Space theory so we will not prove it in this paper. The next theorem is the main result of the paper and presents a construction of a Fractional Brownian Motion with $H > \frac{1}{2}$ using the Haar Wavelets. It is proven that the construction is a Gaussian Process with the correct mean and covariance equation, but continuity is still an open question. A method for proving continuity of the process will be presented as well. It should be noticed that the domain space of the Haar functions introduced in Section 3 is $[0, 1]$, so for convenience we will only consider the parameter space $[0, 1]$ in our construction. If one wants to generalize this construction to all intervals of the form $[0, T]$ then the only requirement is to modify the Haar functions to be be an orthonormal basis on $[0, T]$.

Theorem 5.3. *Take $\{Z_n : n \geq 0\}$ to be an orthonormal basis in H_1, then the random process $\{B^H(t)\}_{t\in[0,1]}$ defined by:*

$$B^H(t) := \sum_{n=0}^{\infty} Z_n \int_0^1 1_{(0,t)}(x) K_H(t, x) P_n(x) dx \tag{5.9}$$

is a Fractional Brownian Motion when $1 > H > \frac{1}{2}$.

Sketch of Proof.

The first fact we need to prove is that this process is, in fact, Gaussian with the proper mean and covariance matrices. If we consider a finite collection $\{B^H_{t_j}\}_{j=0}^m$ with $0 \leq t_1 < t_2 < ... < t_m \leq 1$ then we have

$$E(\exp(i \sum_{j=0}^{m} h_j B^H_{t_j}))$$
$$= \lim_{m\to\infty} \prod_{n=0}^{m} E\left(\exp(iZ_n \sum_{j=1}^{m} h_j \int_0^1 1_{(0,t_j)} K_H(t_j, x) P_n(x) dx) \right)$$
$$= \prod_{n=0}^{\infty} \exp\left(-\frac{1}{2}\left(\sum_{j=1}^{m} h_j \int_0^1 1_{(0,t_j)} K_H(t_j, x) P_n(x) dx \right)^2 \right)$$
$$= \exp\left(-\frac{1}{2} \sum_{j=1}^{m} \sum_{i=1}^{m} h_j h_i \sum_{n=0}^{\infty} \int_0^1 1_{(0,t_j)} K_H(t_j, x) P_n(x) dx \int_0^1 1_{(0,t_i)} K_H(t_i, x) P_n(x) dx \right)$$

where in the 2nd line we exploited the independence of the Z_n's and use the dominated convergence theorem to bring the limit out of the expectation and in line 3 we used the definition of the characteristic function for a Gaussian Random Variable. It follows from the fact that the Haar wavelets are an orthonormal basis in [0,1], Pareseval's identity, and Lemma 4.3 that

$$\begin{aligned}&\sum_{n=0}^{\infty}\int_0^1 1_{(0,t_j)}K_H(t_j,x)P_n(x)dx\int_0^1 1_{(0,t_i)}K_H(t_i,x)P_n(x)dx\\&=\int_0^1 1_{(0,t_j)}K_H(t_j,x)1_{(0,t_i)}K_H(t_i,x)dx\\&=\frac{1}{2}(t_i^{2H}+t_j^{2H}-|t_i-t_j|^{2H})\end{aligned}$$

Thus the above expression becomes:

$$\exp\left(-\frac{1}{2}\sum_{j=1}^{m}\sum_{i=1}^{m}h_jh_i\frac{1}{2}(t_i^{2H}+t_j^{2H}-|t_i-t_j|^{2H})\right).$$

This is the characteristic function of a Gaussian random vector with zero mean covariance function $\frac{1}{2}(t^{2H}+s^{2H}-|t-s|^{2H})$, thus proving that $\{B_{t_j}^H\}$ is Gaussian with the desired mean vector and covariance matrix. This implies that every finite collection of $\{B_t^H\}_{t\in[0,1]}$ is Gaussian with the desired mean vector and covariance matrix, so B_t^H is equal, up to finite dimensional distributions, to a Fractional Brownian Motion. The only fact left to prove is that the paths of B_t^H are continuous with probability one. We were not able to rigorously prove this fact, but we do have a technique for proving it and the way to turn this technique into a formal proof will be addressed at the bottom of this section. For the next few lines please note that an alternate definition for the Haar Wavelets would be

$$P_n(x)=2^{\frac{j}{2}}(1_{(\frac{k}{2^j},\frac{1+2k}{2^{j+1}})}-1_{(\frac{1+2k}{2^{j+1}},\frac{k+1}{2^j})}) \tag{5.10}$$

where, as usual, $n=2^j+k$. From Lemma 5.1 and this alternate definition we can say that:

$$\sum_{n=0}^{\infty}\left|Z_n\int_0^1 1_{(0,t)}(x)K_H(t,x)P_n(x)\right|dx$$
$$\leq X(\omega)\sum_{j=0}^{\infty}\sum_{k=0}^{2^j-1}\log(2^j+k+2)\left|\int_0^1 1_{(0,t)}(x)K_H(t,x)P_{2^j+k}(x)dx\right|$$
$$=X(\omega)\sum_{j=0}^{\infty}\sum_{k=0}^{2^j-1}\log(2^j+k+2)$$
$$2^{\frac{j}{2}}\left|\int_0^t\left(1_{(\frac{k}{2^j},\frac{1+2k}{2^{j+1}})}-1_{(\frac{1+2k}{2^{j+1}},\frac{k+1}{2^j})}\right)K_H(t,x)dx\right|$$

Since K_H is continuous for all $x\in(0,1]$, it is uniformly continuous on $[\rho_j,1]$, where $1>\rho_j>0$. Choose $\rho_j<t$. Thus for $a\in[\rho_j,1]$ and for every $\epsilon>0$ there exists a $\delta>0$ such that whenever $x\in(a-\delta,a+\delta)$ that implies $K_H(t,x)\in(K_H(t,a)-\epsilon,K_H(t,a)+\epsilon)$. Thus if we choose ϵ arbitrarily and a $J\in\mathbb{N}$ such that $2^{-J}<\delta(\epsilon)$ then we have:

$$X(\omega)\sum_{j=J}^{\infty}\sum_{k=0}^{2^j-1}\log(2^j+k+2)\left|2^{\frac{j}{2}}\int_0^t\left(1_{(\frac{k}{2^j},\frac{1+2k}{2^{j+1}})}-1_{(\frac{1+2k}{2^{j+1}},\frac{k+1}{2^j})}\right)K_H(t,x)dx\right|$$
$$=X(\omega)\sum_{j=J}^{\infty}\sum_{k=0}^{2^j-1}\log(2^j+k+2)$$
$$\left|2^{\frac{j}{2}}\int_{\rho_j}^t\left(1_{(\frac{k}{2^j},\frac{1+2k}{2^{j+1}})}-1_{(\frac{1+2k}{2^{j+1}},\frac{k+1}{2^j})}\right)K_H(t,x)dx+\int_0^{\rho_j}1_{(0,t)}K_H(t,x)P_{2^j+k}dx\right|$$
$$\leq X(\omega)\sum_{j=J}^{\infty}\sum_{k=0}^{2^j-1}\log(2^j+k+2)$$
$$\left|2^{\frac{j}{2}}\int_{\rho_j}^t 1_{(\frac{k}{2^j},\frac{1+2k}{2^{j+1}})}(K_H(t,\frac{1+2k}{2^{j+1}})dx+\epsilon)-1_{(\frac{1+2k}{2^{j+1}},\frac{k+1}{2^j})}(K_H(t,\frac{1+2k}{2^{j+1}})-\epsilon)dx+\int_0^{\rho_j}1_{(0,t)}K_H(t,x)P_{2^j+k}dx\right|$$
$$\leq X(\omega)\sum_{j=J}^{\infty}\sum_{k=0}^{2^j-1}\log(2^j+k+2)$$
$$\left|K_H(t,\frac{1+2k}{2^{j+1}})\int_{\rho_j}^1 1_{(0,t)}P_{2^j+k}dx+2\epsilon 2^{\frac{j}{2}}\int_{\rho_j}^1 1_{(0,t)}1_{(\frac{k}{2^j},\frac{1+k}{2^j})}dx+\int_0^{\rho_j}1_{(0,t)}K_H(t,x)P_{2^j+k}dx\right|$$

For $2^j+k\in[2^j,2^{j+1}]$ we can say $\log(2^j+k+2)\leq j+3$ so the above sum can be bounded by

$$X(\omega)\sum_{j=J}^{\infty}(j+3) \tag{5.11}$$
$$\sum_{k=0}^{2^j-1}\left|K_H(t,\frac{1+2k}{2^{j+1}})\int_{\rho_j}^1 1_{(0,t)}P_{2^j+k}dx+2\epsilon\int_{\rho_j}^1 1_{(0,t)}1_{(\frac{k}{2^j},\frac{1+k}{2^j})}dx+\int_0^{\rho_j}1_{(0,t)}K_H(t,x)P_{2^j+k}dx\right|$$

Thanks to the canceling nature of the Haar Wavelets, the first sum converges for any choice of ρ_j. We can choose ρ_j such that the third sum converges. However, there is a problem in the second sum since ϵ is fixed for each j, k, and thus that infinite sum actually diverges. In order to get convergence of the third sum we need for ϵ to depend on j. If we go back to the choice of our ϵ this means that we need to actually make the δ depend on j and derive the appropriate ϵ based on this choice of δ. Since we are working on dyadic intervals the first possible choice of δ_j is fairly natural, namely $\delta_j = 2^{-j}$. With this choice of δ_j it also seems reasonable that the choice of ϵ_j would be $\epsilon_j = \max_{k\leq 2^j} \max_{x\in(\frac{k}{2^j},\frac{k+1}{2^j})} \{K_H(t,x)\} - \min_{k\leq 2^j} \min_{x\in(\frac{k}{2^j},\frac{k+1}{2^j})} \{K_H(t,x)\}$. We can find epsilon by bounding the derivative, but since the derivative becomes infinite at zero this means that the convergence of the middle sum will depend on the choice of ρ_j. We could fix ρ so it does not depend on j and then the middle sum will behave nicely, but then one must use the cancellation property of the Haar Wavelets to bound the final sum and we have not discovered how to take advantage of this yet. This completes the sketch of the proof.

With a little inspection one can notice that the above representation can be extended to the $H = \frac{1}{2}$(Brownian Motion) case simply by setting $K_{\frac{1}{2}}(t,s) = 1$. If we do this and follow the steps in Lemma 4.4 in the exact same manner we get $\int_0^T 1_{(0,s)}(x)1_{(0,t)}(x)dx = min\{s,t\}$, which is the covariance function of a Brownian Motion. Thus by using $K_{\frac{1}{2}} = 1$ in (5.9) we have a natural extension of our construction to the Brownian Motion case. Continuity has also been established for the $H = \frac{1}{2}$ case, so we can say that this representation truly works for $H = \frac{1}{2}$.

6 Conclusion

The core of the construction presented in the previous section is the function K_H. For $H \geq \frac{1}{2}$ we have now shown that (5.9) is a good candidate for being a Fractional Brownian Motion, but we would like to extend this construction to

$H < \dfrac{1}{2}$ if possible. While we do not present this extension in this thesis, the intuition that provided the above construction might be implemented to come up with a construction for $H < \dfrac{1}{2}$.

As mentioned in Section 4, the intuition behind using the K_H for the representation (5.9), which comes from the theory of Reproducing Kernel Hilbert Spaces, is that K_H satisfies the equation

$$R_H = \int_{[0,T]} 1_{(0,s)}(v) 1_{(0,t)}(v) K_H(s,v) K_H(t,v) dv \tag{6.12}$$

for $H > \dfrac{1}{2}$. Analogously, if one could find a K_H that would satisfy the above equation for $H < \dfrac{1}{2}$ then the resulting equation might help to extend (5.9) to a well defined B_H for all $H \in (0,1)$.

It turns out that

$$K_H(t,s) = b_H \left[\left(\frac{t}{s}\right)^{H-\frac{1}{2}} - \left(H - \frac{1}{2}\right) s^{\frac{1}{2}-H} \int_s^t (u-s)^{H-\frac{1}{2}} u^{H-\frac{3}{2}} du \right] \tag{6.13}$$

with $b_H = \sqrt{\dfrac{2H}{(1-2H)\beta(1-2H, H+\frac{1}{2})}}$ satisfies the equation (6.12) for $H < \dfrac{1}{2}$. For a full proof of this fact one can look at [DU99]. While much remains to be justified, we conjecture that Theorem 5.3 might be extended to $H < \dfrac{1}{2}$ by using (6.13).

References

[Ste01] Steele, J. Michael. Stochastic calculus and financial applications. Vol. 45. Springer Science, Business Media, 2001.

[Bia08] Biagini, Francesca, et al. Stochastic calculus for fractional Brownian motion and applications. Springer Science, Business Media, 2008.

[Mis08] Mishura, Yuliya. Stochastic Calculus for Fractional Brownian Motion and Related Processes. Springer, 2008.

[Mal09] Mallat, Stéphane. A Wavelet Tour of Signal Processing. Elsevier/Academic Press, Amsterdam, 2009.

[TM93] Hida, Takeyuki and Hitsuda, Masuyuki. Gaussian Processes. American mathematical Society, Providence, RI, 1993.

[EHK03] Bayraktar, Erhan and Poor, H. Vincent and Sircar, K. Ronnie. Estimating the Fractal Dimension of the S&P 500 Index using Wavelet Analysis. 2003.

[RH88] Barton, Richard R. and Poor, H. Vincent. Signal Detection in Fractional Gaussian Noise. IEEE Transactions on Information Theory, Vol. 34, no. 5. 1988.

[DU99] Decreusefond, L. and Ustunel, A.S. Stochastic Analysis of the fractional Brownian motion. Potential Analysis 10, 177-214, 1999.

[ID92] Daubechies, Ingrid. Ten lectures on wavelets. Vol. 61. Philadelphia: Society for industrial and applied mathematics, 1992.