TECHNICAL RESEARCH REPORT

SYSTEMS RESEARCH CENTER

Supported by the National Science Foundation Engineering Research Center Program (NSFD CD 8803012), the University of Maryland, Harvard University, and Industry

Reconstruction of Stochastic Processes Using Frames

by J.T. Gillis

TR 91-14

Reconstruction of Stochastic Processes Using Frames

J. T. Gillis*

February 12, 1991

Abstract

This note discusses sampling in a general context and shows that the (dual) frame reconstruction formula holds for stochastic processes, in quadratic mean. Specifically we show that if the covariance can be reconstructed using frames then the sample path can also be reconstructed. The application of the result for the generation of approximate sample paths for simulation is discussed. Ergodic properties of the approximate estimators are also investigated.

1 Introduction

In this paper we shall discuss the construction and reconstruction of stochastic processes using frames and wavelets. We are motivated by the need to generate paths in order to solve (simulate) equations involving stochastic processes. The fundamental principle is (re)construction of the sample path from samples or some appropriate generalization of sampling such as linear functionals of the sample path. The paper opens with some general remarks on sampling theory. Section two is a discussion of some of the potential problems of using frames for sequential approximation. The third section of the paper discusses a generalization of the stochastic version of the Shannon—Whittaker theorem to frames. Simulation of sample paths using this construction is discussed, including some remarks on ergodic properties of the representation.

The reconstruction of a function from data about the function has a long and important history in applied mathematics. Let $f(\cdot)$ denote the function to be reconstructed, $s(\cdot)$ be a known function, and define the translation operator $(T_{t_n}s)(\cdot) \mapsto s(\cdot - t_n)$. Let $c(f)$ denote a functional of the signal to be reconstructed and $F(f)$ be some function(al) of the unknown function which we desire to know, e.g. f is a stochastic process and F the covariance function.

*Currently at the Systems Research Center, University of Maryland, College Park, MD. on leave from the The Aerospace Corporation, Los Angeles, CA.

Signals used as interpolation functions are denoted as $s_n(t)$ or $(T_{t_n}s)(t)$. There are at least five types of reconstruction results that can be achieved with these ideas:

I $f(t) = \sum_n f(t_n)T_{t_n}s(t)$

II $f(t) = \sum_n f(t_n)s_n(t)$

III $f(t) = \sum_n c_n(f)T_{t_n}s(t)$

IV $f(t) = \sum_n c_n(f)s_n(t)$

V $F(f)(t) = \sum_n c_n(f)s_n(t)$

Note that if $c_n(f) = \int \delta(\tau - t_n)f(\tau)d\tau$, then type I is a special case of type III. In fact, the order is of increasing generality: type I $\subset$ type II $\subset$ type III $\subset$ type IV $\subset$ type V. As a subjective matter they are numbered in decreasing order of importance, at least when special cases are considered. Obviously, necessary and sufficient conditions on f, c, and s for theorems of type IV would be quite a feat. Such conditions would have to include all of the known theory of Fourier series!

In addition, the function being reconstructed can be either deterministic or stochastic. If the function is stochastic then there are (at least) three types of reconstruction that can occur: spectral (or second order mean and covariance), quadratic mean, and path-wise (for almost all paths). Further, the data can be generated in a deterministic or stochastic way (e.g. sampling on a lattice or by sampling at the jumps of a Poisson process). Deterministic sampling can be regular, e.g. based on a lattice, or irregular. Stochastic sampling of gathering stochastic data can be dependent, although we are not aware of any such results, or independent, for which there are several results[1]. If reconstruction is such that the function (sample path) can be uniquely identified within a class, then the reconstruction is said to be *alias–free*. If the function is only determined up to a subclass, then the members of the subclass are known as the *aliases* of the function. A sampling scheme may be alias–free in one sense and not in another. The best example of this is that there is a sampling scheme, based on a counting measure, which is shown alias–free by Masry[2]. It is shown that the counting measure allows the power spectrum of the process to be reconstructed. The same measure is not alias–free for reconstruction of the spectrum by Shapiro and Silverman's criteria [3]. Both theorems are correct; what *is* different is the way in which data is collected and used to determine if the sampling scheme is alias–free. Shapiro and Silverman use the *correlation function of the samples* and Masry uses the *law of the compound process*. Shaipro and Silverman's result includes the classical Nyquist theorem [4].

[1] Of course, there are reconstruction theorems other than the types mentioned for stochastic sampling. In particular, the Law of the samples may determine the Law of the sampled processes, etc.[1]

Further, Masry [5] gives and example of a sampling scheme which is alias–free in the sense of Shapiro and Silverman which cannot lead to path based estimates with probability one[2].

The famous result of Whittaker and Shannon for the reconstruction of Paley–Wiener (bandlimited) functions from samples taken on an appropriate lattice has a direct stochastic analog (see Section 6).

The following table gives (an incomplete) summary of typical theorems.

Function Data	Lattice	Sampling Measure Irregular	Stochastic
$f(t)$, Det. $f(t_k)$	I: Shannon[6]	II, III: Levinson[7] Benedetto–Heller[8]	Shaipro–Silverman[3] Beutler[9], Masry[2]
$R(\tau)$, Cov. X_{t_k}	V: Masry[2] Shapiro–Silverman[3]	?	V: Papoulis[10]
$P(\omega)$, Law. $Prob_{X_{t_k}}$	?	?	Lewis[11], Karr[1]
X_t, Q. M. X_{t_k}	I: Balakrishnan [12]	II: Benedetto [13] IV: This paper	II: Gillis (Asym.) [4]
X_t, Path	I: Houdré[14]	?	?

The main theorem of this paper is roughly this: if one can reconstruct the covariance of a processes using frames, then one can reconstruct the sample path in quadratic mean using frames. The result is interesting because of its usefulness in simulations, which is discussed in some detail in the paper. It is also interesting because the "meta theorem" that says if one can reconstruct the covariance then one can reconstruct the sample path *is false* in general. This paper provides a partial answer to when the theorem *is* true. The counterexample to the meta theorem has the following brief summary[4]: Poisson sampling is the best sampling scheme in a wide class of stochastic sampling measures (stationary point processes) uncorrelated with the sampled processes. Poisson sampling is not alias–free for reconstruction of quadratic mean continuous processes (using type II estimators). The process is recovered as the rate increases (asymptotically).

It had been previously shown that Poisson sampling is alias–free for the reconstruction of the covariance of most any process (an informal discussion is found in Papoulis [10]); however

[2]The sampling scheme is based on $t_n = n\tau$, where τ is exponentially distributed.

the reconstruction is based on a nonlinear function of the data. Other methods of recovering the covariance depend on the sense that aliasing is meant (Masry, Shapiro–Silverman, etc). Our current view on this is that Poisson sampling gives the magnitude, but not the phase, of the spectral process (the processes' Fourier transform).

For the simulation of integral and differential equations with random components, two senses of solution are important: solutions in law (or second order solutions) and path solutions. If the simulation is of the Monte Carlo type, then the solution must have the correct distribution; however, it need not have any path properties. For instance, in dispersion studies of launch vehicle trajectories, it is only the second order properties (mean and variance) that are of interest. Since the actual trajectory that will be flown depends on the winds at the time of launch, only statistical properties of the winds–aloft and the responses to them are important for simulation. The second use for a simulation is to reconstruct events, such as the exact behavior of a system. In this case the path behavior of the system is important. For example, if one is trying to understand (and correct) anomalous behavior of a spacecraft, one of the uses of a simulation is to verify that all of the pertinent phenomenology has been modeled. Verification of the model is demonstrated by reproducing the anomaly (i.e. the path). When dealing with path solutions, several types of convergence are possible; two distinguished types are almost sure and quadratic mean. While one would prefer to have almost sure convergence, it is routine to settle for quadratic mean convergence in engineering applications, as it is much easier to establish.

2 Frame Based Estimation

The use of frames to reconstruct functions dates to the paper of Duffin and Schaeffer [15]. Given a Hilbert Space H, a set of vectors $\{\phi_j\} \subset H$, is called a *frame* if

$$\exists\, A, B \text{ with } 0 < A \leq B < \infty \text{ s.t. } A||f||^2 \leq \sum_j |< \phi_j,\ f >|^2 \leq B||f||^2 \qquad \forall f \in H$$

The constants A and B are called the frame bounds. If $\frac{B}{A}$ is close to one, then the dual frame representation (see below) converges relatively fast. All of the facts we will use about frames are contained in Ingrid Daubechies' paper [16]. Given a frame, it can be shown that two reconstruction formulas for a function $f \in H$ are valid:

1. Frame Representation: $f(t) = \sum_j < \mathcal{T}^{-1}\phi_j,\ f > \phi_j$
2. Dual Frame Representation: $f(t) = \sum_j < \phi_j,\ f > \mathcal{T}^{-1}\phi_j$

where operator $\mathcal{T}^{-1}$ acts on the frame elements to produce a "dual frame." The operator $\mathcal{T}$ is given explicitly as $\mathcal{T} = T^*T$, where $Tf \mapsto \{< f,\ \phi_j >\}_j$, which is a bounded invertible linear operator. The set of vectors $\{\mathcal{T}^{-1}\phi_j\}_j$ are a frame in H also, with frame constants $\frac{1}{B}$ and $\frac{1}{A}$. The related operator $G = TT^*$ is given by the matrix of values $\{< \phi_i,\ \phi_j >\}$ and the spectrum of G and $\mathcal{T}$ are the same.

This represents type IV signal reconstruction. Further, the lower bound is know in the "pre–wavelets" literature as the minimum modulus and it is greater than zero if and only if the dimension of the null space of the operator T is zero and the range of T is closed [17][18][19]. We note that the stability theory for the operator T, and hence $\mathcal{T}^{-1}$ is well developed. This is clear once one notes that T is a bounded semi–Fredholm operator[17] with $ker(T) = \{0\}$. Every bounded injective semi–Fredholm operator from L_2 into l_2 is associated with a frame (use coordinate projection and then the Riesz representation theorem). The lower frame bound is the minimum modulus, denote as γ_T. The following theorem is far from the last word on the subject, but it gives one confidence that the frame property is robust.

Theorem 1 (First Stability Theorem [17]) *Let $T : X \to Y$ be a closed linear operator and semi–Fredholm with minimum modulus γ_T. Let A be a T bounded operator (i.e. $||Au|| \leq a||u|| + b||Tu||$ $u \in D(T)$), with $a < (1-b)\gamma$. Then $S = A + T$ is semi–Fredholm and $dim(ker(S)) \leq dim(ker(T))$, $codim(Ra(S)) \leq codim(Ra(T))$, and $index(S) = index(T)$.*

Corollary 2 *Let T_{ϕ_j} be the operator $Tf \mapsto \{< \phi_j,\ f >\}$ and let $\{\phi_j\}_j$ be a frame. Let S be the operator $Sf \mapsto \{< \psi_j,\ f >\}$. If $||S|| \leq \gamma_{T_{\phi_j}}$, then $\{\phi_j + \psi_j\}_j$ is a frame.*

As previously mentioned bounded linear operators $S : L_2 \to l_2$, are all of the form $Sf \mapsto \{< \psi_j,\ f >\}_j$. If T is associated with a frame and S a linear bounded operator, then $T + \frac{\gamma_T}{2||S||}S$ is associated with a frame (via projection and the Riesz map). In addition, one could investigate new wavelet bases by this method, e.g. if the action of a group G on a function, ϕ generates a frame, when does the action of the group on $\phi + \psi$ generate a frame?

The theory of semi–Fredholm operators also allows one to examine "functions which are frames for their span" type of questions, by using the concept of the reduced minimum modulus[17], and by looking at $L_2/ker(T_{\phi_j})$ and $Ra(T_{\phi_j})$.

One can give geometrical interpretations of the minimum modulus in terms of the graph of the operator and in terms of the spectrum of the operator [18][20].

Since we can represent $L_2(I\!R)$ functions using frames, it is natural to try to represent stochastic processes in this manner. In examining the properties of a stochastic process, with mean and covariance in L_2, we see that the sample paths are not necessarily L_2 functions [6], so the representation is not automatic.

That is, we shall examine quadratic mean path reconstruction and give conditions under which it can be achieved for stochastic processes using frames. In analogy with the cardinal sampling case, it will turn out the reconstructability of the covariance is sufficient for quadratic mean reconstruction of the path. First we shall examine several issues that arise when one uses frames in any sequential approximation scheme.

3 Sequential Approximation and Frames

In this section, we give several observations on the use of frames for sequential approximation in the solution of integral and differential equations as well as in nondynamic simulations.

The first observation is that any finite set of vectors is a frame for its span.

Theorem 3 *Let $X = \{\phi_j\}_j^N$ be a finite set of vectors in H, a Hilbert space. Let $H_\phi = Span\{\phi_j\}$; then X is a frame for H_ϕ.*

Proof.

Let $f(x) = \sum_j |< x,\ \phi_j >|^2$, which is a continuous function, with domain $H = Span\{\phi_j\}$. Also, $||x||^2 f(\frac{x}{||x||}) = f(x)$. Clearly $f(x) \geq 0$. The function attains its maximum and minimum on the compact set $\{x|||x|| = 1\}$; let this maximum be B, and the minimum be A. Let x_* be the minimizing element, i.e. $f(x_*) = A$ and $||x_*|| = 1$. Hence

$$A||x||^2 \leq \sum_j |< x,\ \phi_j >|^2 \leq B||x||^2;$$

However, it remains to show that $A > 0$. If $A = 0$, then $x_* \perp \phi_j\ \forall\ j$, therefore $||x_*|| = 0$. This is a contradiction, as we assumed that $||x_*|| = 1$. Hence $A > 0$, and A and B are the upper and lower frame bounds. ∎

However, not every collection of vectors is a frame for its span. Let e_j be the usual basis for l_2, that is: $e_1 = (1, 0, \cdots)$, $e_2 = (0, 1, 0, cdots)$, etc. Consider $\{\frac{1}{n}e_j\} \subset l_2$, which spans l_2, and is not a frame. Nor is every subset of a frame a frame. Consider the frame given by:

$$\phi_j = \begin{cases} e_{\frac{j}{2}} & j\ even \\ \frac{1}{j}e_{\frac{j+1}{2}} & j\ odd \end{cases}$$

The set is easily shown to be a frame for l_2 with frame bounds $A = 1$, $B = 2$. The "even index" subset is an orthonormal basis, and therefore a frame. The "odd index" subset is clearly not a frame. We need a stronger concept , namely:

Definition 4 *A uniformly approximating sequence (UAS) is a sequence of nested subsets of a frame which satisfy:*

1. *Each subset is a frame for its span.*
2. *The frame constants are uniformly bounded. In particular the lower frame bounds are bounded away from zero.*

The upper frame bound holds for any subset of a frame, so the main concern is always with the lower frame bound.

In the last example, let $H_n = Span\{\phi_j\}_{j=1}^n$ and $F_n = \{\phi_j\}_{j=1}^n$. All of the sets F_n, are frames for H_n, respectively. The sequence F_{2n} is a UAS for l_2 it is sequence of orthonormal basis. In contrast the sequence F_{2n+1} is not a UAS, since the lower frame bound for F_{2n+1} in H_{2n+1} is $\frac{1}{2n+1}$.

The importance of this definition comes in the following way: say we are using a scheme, such as a the Galerkin method, to solve an equation in a function space. If the solution is going to zero in a UAS, then it is going to zero in the strong sense. If the set is not a UAS then it could be going to zero in the weak sense.

The sum of two frames is not necessarily a frame, as the following counterexample shows (developed with Y. Pati).

Example 5 *In l_2, let $\psi_j = e_{3j}$, and $\phi_j = \sqrt{1-\frac{1}{j}}e_{3j} + \sqrt{\frac{1}{j}}e_{3j+1}$. The sets $X = \{\psi_j\}_j$ and $Y = \{\phi_j\}_j$ are frames; in fact they are orthonormal sets. The combined span is the set $Span\{\psi_j, \phi_j\} = Span\{e_{3j}, e_{3j+1}\}$. The following calculation: $\frac{1}{k} = \sum_j |< e_{3k+1},\ \phi_j >|^2 + |< e_{3k+1},\ \psi_j >|^2$, shows that this set is not a frame for the joint span.*

When combining frames or using subsets of frames, one must be careful to choose the way in which the approximation proceeds.

Clearly some frames are UAS's; every appropriate subdivision of an orthonormal set is a UAS. Thus, a multiresolution analysis (MRA) [21] is a UAS. Current work with Y. Pati and P. S. Krishnaprasad gives a sufficient condition for some wavelet bases to be a UAS in the same way that an orthonormal set is a UAS. Better understanding of this issue is quite important in the interpretation of solutions based on refinement algorithms (e.g. finite element mesh refinement algorithms, etc.).

The basic circle of ideas is taken from classical ideas about least squares approximation. Recall that the frame operator $\mathcal{T}$ has the same spectrum as G, which is the Gram matrix, $\{G\}_{ij} = \{< \phi_i,\ \phi_j >\}$. For a finite set of vectors, we have the Gram matrix $G_n = \{G_n\}_{i,j=1\cdots n} =$

$\{<\phi_i,\ \phi_j>\}_{i,j=1\cdots n}$. A quick calculation shows that the maximum and minimum eigenvalues of G are the maximum and minimum singular values of T. Hence the maximum and minimum singular values of G are the frame bounds for T. Further, the singular values allow interpretation as $\cos(\theta_j)$, where theta is the principal angle between the vector ϕ_j and $\{\phi_i\}_{\{i=1,\cdots j-1\}}$ [22]. It is sufficient for a set to be a UAS that the smallest singular value of T be bounded from below. Hence, a frame will be a UAS if there is a minimum "angle" between the vectors. If the frame elements are all of unit length, we are requiring that $1 > \tilde{A} = sup_{i\neq j}|<\phi_i,\ \phi_j>|$. This condition would seem sufficient for the set to be a UAS for its span, when arranged appropriately (nested etc). Because wavelets are generated by a single function and a group action, one can check to see that there is a minimum angle between the subspaces directly. This argument is sufficient for the finite dimensional case. Strictly as stated, this argument does not allow passage through the limit, because the spectrum of the operator could become more complicated than the point spectrum. In functional analysis there are two simple cases that allow an operator to have a pure point spectrum – the operator being compact or the operator having compact resolvent. Since the Gram matrix for an orthonormal set is the identity on l_2, the Gram matrix does not have to be compact; one can show using spectral considerations that it cannot be compact. If the Gram matrix has compact resolvent then it has a sequence of eigenvalues that tend to infinity [23], which is precluded by the upper frame bound. If one examines the frame given by $\{\frac{1}{\sqrt{1+n}}e_n\}$ then the 1 is seen to be an accumulation point of the spectrum, of both T and G.

The sufficiency arguments developed by Y. Pati make these ideas precise; indeed his arguments are more general than this argument . They will be reported elsewhere.

4 The Wiener-Hopf Equation

In this section, we examine the simulation of the sample path of a random processes. The two basic problems are the simulation of a random process with the appropriate properties and the reconstruction of the sample path (in quadratic mean). We shall show that there is a correspondence principle: if the covariance is reconstructable then the sample path is also.

If we attempt to formally write the frame and dual frame representations, we find integrals of the form: $\int X_t f(t)\, dt$. It can be shown that this integral exists as a quadratic mean integral *if and only if* the Riemann integral

$$\int\int f(t)\overline{f(s)}R(t-s)\ dtds$$

exists [6]. The integral $\int f(t)X_t\, dt$ can be defined as a Lebesgue integral for almost all paths if the process is measurable and

$$\int |f(t)|E|X_t|\, dt < \infty$$

(op. cit.).

Let $\{X_t\}$ be a wide sense stationary stochastic process, zero mean and covariance function $R \in L_2$. We define an estimator of X_t (at the point t) given the data $\{< \mathcal{T}^{-1}\phi_j,\ X >\}_j$ or $\{< \phi_j,\ X >\}_j$ as

$$\widehat{X_t} = \sum_j < \phi_j,\ X_t > \mathcal{T}^{-1}\phi_j(t)$$

which is called the dual frame estimator and

$$\widehat{X_t} = \sum_j < \mathcal{T}^{-1}\phi_j,\ X_t > \phi_j(t)$$

which is known as the frame estimator. Since the formulae are symmetric with respect to the frame and the dual frame, and the "double dual" frame is the original frame, henceforth we shall write both estimators as:

$$\widehat{X_t} = \sum_j < \phi_j,\ X_t > \psi_j(t).$$

The understanding is that $\{\phi_j\}$ and $\{\psi_j\}$ are associated as dual frames.

Since $R \in L_2$, we have the frame and dual frame representation for R:

$$R(t) = \sum_j < \phi_j,\ R > \psi_j(t),$$

which will be used repeatedly in subsequent calculations.

Given the data $\{< \phi_j,\ X_t >\}_j$, in the Hilbert space H^X generated by the random variables X_t, a necessary condition[24] for minimum variance linear estimation is that the estimation error be orthogonal to the data — i.e. $\forall s$:

$$0 = E\{(X_t - \widehat{X_t})X_s\}$$

Verify:

$$\begin{aligned}
E\{(X_t - \widehat{X_t})X_s\} &= R(t-s) - \sum_j \int \phi_j(\tau)\ E\{X_\tau X_s\}\ d\tau\ \psi_j(t) \\
&= R(t-s) - \sum_j \int \phi_j(\tau) R(\tau - s)\ d\tau\ \psi_j(t) \\
&= R(t-s) - R(t-s) = 0,
\end{aligned}$$

as required.

The other quantity of interest is the estimation error:

$$\begin{aligned}
\mathcal{E}_t &= E\{(X_t - \widehat{X}_t)^2\} \\
&= E\{(X_t - \widehat{X}_t)X_t\} - E\{(X_t - \widehat{X}_t)\widehat{X}_t\} \\
&= -E\{(X_t - \widehat{X}_t)\widehat{X}_t\} \\
&= E\{\widehat{X}_t\widehat{X}_t\} - \sum_j \int \phi_j(\tau)\ E\{X_\tau X_t\}\ d\tau\ \psi_j(t) \\
&= E\{\widehat{X}_t\widehat{X}_t\} - \sum_j \int \phi_j(\tau)\ R(\tau - t)\ d\tau\ \psi_j(t) \\
&= E\{\widehat{X}_t\widehat{X}_t\} - R(0) \\
&= E\left\{\left(\sum_j \int \phi_j(\tau)\ X_\tau\ d\tau\ \psi_j(t)\right)\left(\sum_k \int \phi_k(\sigma)\ X_\sigma\ d\sigma\ \psi_k(t)\right)\right\} \\
&\quad -R(0) \\
&= \sum_j \sum_k \int\int \phi_j(\tau)\phi_k(\sigma)\ R(\tau - \sigma)\ d\tau d\sigma\ \psi_j(t)\psi_k(t) - R(0) \\
&= \sum_j \int \sum_k < R(\cdot - \sigma),\ \phi_k > \psi_k(t)\ \phi_j(\sigma)\ d\sigma\ \psi_j(t) - R(0) \\
&= \sum_j \int \phi_j(\sigma)R(t - \sigma)\ d\sigma\ \psi_k(t) - R(0) \\
&= R(0) - R(0) = 0
\end{aligned}$$

so that the reconstruction is perfect (in quadratic mean). Hence we have the following theorem:

Theorem 6 (Correspondence Principle) *Let X_t be a wide sense stationary stochastic process, with covariance R. Let $\{\phi_j\}_j$ and $\{\psi_j\}_j$ be dual frames such that:*

$$R(t) = \sum_j < R,\ \phi_j > \psi_j(t)$$

Then

$$X_t = \sum_j < X,\ \phi_j > \psi_j(t)$$

in quadratic mean.

Note that the estimator minimizes almost any functional of the error (e.g. $E\{\int_{-T}^{T}(X_t - \widehat{X}_t)^2\ dt\} = E\{\left|\left|X - \widehat{X}\right|\right|^2_{L_2(-T,T)}\}$).

5 Stochastic Frames

Thus far we have apparently dealt with deterministic sampling; if we assume that we have a "stochastic frame," i.e. a set of vectors in H^X such that the representation:

$$\breve{R}(t) = \sum_j < \phi_j,\ R > \psi_j$$

holds in such a way that $E_\tau\{(f - \breve{f})^2\} = 0$, then once again a correspondence principle will allow us to reconstruct the sample path of the process using the "stochastic frame." Here τ is a random parameter (independent of X_t, when the function f is taken as a stochastic process), involved in the development of the frame (e.g. the frame is generated by translations of a fixed function and the translation are allowed to have "jitter" in the sampling time). Then the above arguments can be modified by factoring E as $E_\tau E_X$ and applying first E_X and then using E_τ, we have the representation:

$$\breve{X}_t = \sum_j < \phi_j,\ X > \psi_j,$$

such that $E\{(X_t - \breve{X}_t)^2\} = 0$. As usual, ϕ_j and ψ_j are dual frames and equality is in quadratic mean.

Conditions under which a set of functions can be developed as a "stochastic frame" are under investigation by Prof. John Benedetto and his students, at least for the case of wavelets [3].

6 Example: Shannon–Balakrishnan

In this section we shall show that the correspondence principle implies the Stochastic version of the Whittaker–Shannon theorem directly from the classical Whittaker–Shannon theorem.

Consider the Paley–Wiener space $PW_\Omega = \{f \in L_2\ s.t.\ supp(\hat{f}) \subset (-\Omega, \Omega)\}$; recall that there is a (bounded) reproducing functional (e.g. a d_Ω s.t. $(d_\Omega * f)(t) = f(t)$) on this space given by:

$$d_\Omega(t) = \frac{\sin 2\Omega\pi(t)}{2\Omega\pi(t)}$$

Note that $d_\Omega * f(t) = < T_t d_\Omega,\ f >$, and that the frame operator is the identity for an orthonormal basis, and so the frame is self dual.

The now classical theorem of Whittaker–Shannon is:

[3]The author would like to thank Prof. Benedetto for his scholarly exposition of frames and wavelets given in his Wavelets Seminar at the University of Maryland, Fall 1990.

Theorem 7 (Whittaker-Shannon [25]) *If $f \in PW_\Omega$, then*

$$\begin{aligned} f(t) &= \sum_{n \in \mathbb{N}} f(\frac{n}{2\Omega}) \frac{\sin 2\Omega\pi(t - \frac{n}{2\Omega})}{2\Omega\pi(t - \frac{n}{2\Omega})} \\ &= \sum_n < f,\ T_{\frac{n}{2\Omega}} d_\Omega > T_{\frac{n}{2\Omega}} d_\Omega(t) \end{aligned}$$

From this result one can conclude:

Theorem 8 (Balakrishnan[12]) *If X_t is Wide Sense Stationary with $E\{X_t X_s\} = R(t-s) \in PW_\Omega$, then*

$$\begin{aligned} X_t &= \sum_n < X,\ T_{\frac{n}{2\Omega}} d_\Omega > T_{\frac{n}{2\Omega}} d_\Omega(t) \\ &= \sum_{n \in \mathbb{N}} X_{(\frac{n}{2\Omega})} \frac{\sin 2\Omega\pi(t - \frac{n}{2\Omega})}{2\Omega\pi(t - \frac{n}{2\Omega})} \end{aligned}$$

The only statment that needs to be proved is $X_t = < X,\ T_t d_\Omega >$, and this is clear from the spectral representation of the processes [6] and Fubini's Theorem.

7 Ergodic Properties

Here we are thinking of second order processes as actually being Gaussian, and working with the second order properties of the process.

The spectral representation of a stochastic process indicates that

$$X_t = \int e^{i2\pi\ t\omega}\ \hat{X}(d\omega)$$

so that one might be tempted to approximate the stochastic process by:

$$\begin{aligned} X_t &\approx \sum_j e^{i2\pi\ t\omega_j} \int_{A_j} \hat{X}(d\omega) \\ &\approx \sum_j \cos(2\pi t\omega_j)\ Y_j + \sin(2\pi t\omega_j)\ Z_j \end{aligned}$$

where Y_j and Z_j are uncorrelated random variables, and $\omega_j \in A_j$. While one can show that such approximations converge as $N \to \infty$ and $sup|A_j| \to 0$ (at least in some cases), the approximation is not ergodic. This is a consequence of Maruyama's Theorem [26], which states that a Gaussian process is ergodic *if and only if* the spectral–distribution function is continuous.

We take the approximation of X_t as:

$$X_t \approx \check{X}_t = \sum_{j=-N}^{N} < X,\ \phi_j > \psi_j(t).$$

Then

$$E\check{X}_t = \sum_{j=-N}^{N} \int E\{X_\tau\}\phi_j(\tau)\, d\tau\ \psi_j(t) = 0$$

and

$$\begin{aligned}
E\check{X}_s\check{X}_t &= \sum_{k=-N}^{N}\sum_{j=-N}^{N} \int\int E\{X_\tau X_\sigma\}\phi_k(\sigma)\phi_j(\tau)\, d\tau d\sigma\ \psi_k(t)\psi_j(s) \\
&= \sum_{k=-N}^{N}\sum_{j=-N}^{N} \int\int R(\tau-\sigma)\phi_k(\sigma)\phi_j(\tau)\, d\tau d\sigma\ \psi_k(t)\psi_j(t) \\
&\approx \sum_{k=-N}^{N} \int R(t-\sigma)\phi_k(\sigma)\, d\sigma\ \psi_j(s) \\
&\approx R(t-s)
\end{aligned}$$

One important aspect of the frame representation is that it can be used to develop approximations of the stochastic process in such a way that the approximate sample paths are ergodic. Of foremost importance is the class of frames know as wavelets. It can be shown (again the paper of I. Daubechies [16] is an excellent source for the ideas that are used here) that a wavelet frame is a frame generated by a single function (analyzing–wavelet) via the interaction of two group operations. If translations and dilations of the analyzing–wavelet generate the frame, then it is referred to as an affine wavelet. If the frame is generated by translation in both the time domain and in the frequency domain of the analyzing wavelet, then the frame is known as a Weyl–Heisenberg wavelet.

For Gaussian random processes the requirement that the approximate sample function be ergodic is easy to check; it is that the covariance $R_N(t,s) = E\{\check{X}_t\check{X}_s\}$ have a continuous spectral distribution function[26].

$$\begin{aligned}
\widehat{R_N}(\omega) &= \int \sum_{k=-N}^{N}\sum_{j=-N}^{N} \int\int R(\tau-\sigma)\phi_k(\sigma)\phi_j(\tau)\, d\tau d\sigma\ \ \psi_k(t)\psi_j(t)\ \ e^{i2\pi t\omega} dt \\
&= \sum_{k=-N}^{N}\sum_{j=-N}^{N} c_{k,j}(R) \int \psi_k(t)\psi_j(t)\, e^{i2\pi t\omega}\, dt \\
&= \sum_{k=-N}^{N}\sum_{j=-N}^{N} c_{k,j}(R)\, (\hat{\psi}_k * \hat{\psi}_j)(\omega)
\end{aligned}$$

Hence the ergodicity of the approximate processes depends on the relationships of the frame elements to each other. Both affine wavelets and Weyl-Heisenberg wavelets can satisfy this additional constraint. Further the use of wavelets allows control of the accuracy of the approximation that has time and frequency localization. This allows for the adaptive generation of a sample path of a stochastic process, including the possibility of "rollback" in a simulation. If at some point the approximation of the sample path has been too coarse, then return to some time when the simulation was correct and proceed forward (on the same sample path), from that point with higher fidelity; perhaps even relaxing the precision of the approximation after a disruption has occurred. Simulations of systems that contain significant noise contributions and which have ' 'switching elements[4]" are currently fairly expensive to develop, as the only way to find the switching time is to overshoot it and then backup (i.e. rollback). Simply calling a random number generator to develop a sample path may generate an inconsistent sample path, additionally the sample variance is then dependent on the integration step size. The sample paths of the noise must then be generated by a method such as Mercer's expansion (also called the Karhunen–Loéve expansion [6]), which has no particular time–frequency localization properties and is peculiar to each process.

In order to generate an approximate sample path one must generate random variables (or more correctly pseudo–random variables) with the correct properties:

$$E< X,\ \phi_j > = \int EX_t\phi_t(t)\ dt = 0$$

and

$$\begin{aligned}
&E< X,\ \phi_j >< X,\ \phi_k > \\
&= \int\int EX_tX_s\ \phi_j(t)\phi_k(s)\ dtds \\
&= \int\int R(t-s)\ \phi_j(t)\phi_k(s)\ dtds \\
&= \int R(t')\int \phi_j(t'+s)\phi_k(s)\ dsdt' \\
&= -\int R(t)\ (\phi_j * \phi_k)(t)\ dt
\end{aligned} \tag{1}$$

While equation 1 may not be easy to solve, one can generate the functions $\phi_j * \phi_k$ in advance and use factorization techniques to generate the appropriate coefficients. Further, if the frame is a wavelet, then the convolution can always be transformed by a change of variables to be of the form $\phi_0 * G_{g_j,g_k}\phi_0$, where G_{g_j,g_k} represents the group action. Then one generates random variables according to:

$$R_N = E< X,\ \phi_j >< X,\ \phi_k >,$$

which is a (finite?) matrix; it is non–negative definite and therefore can be factored as: $R_N =$

[4]That is elements with jump discontinuities

L^*L. Let $\{x_k\}$ be i.i.d., N(0,1) random variables. Then $y_k = \sum_j L_{k,j} x_j$ are random variables with the desired correlation structure.

8 Summary

In summary, the theory of frames and especially wavelets appears to meet the required criterion for use in the simulation of systems with random components. Currently we are developing a simulation capability to explore the possibilities for the generation and reconstruction of random processes (and random fields) using wavelets. While the frame–wavelet concept does not offer a panacea for the computation and reconstruction of random processes, it offers some substantial benefits. If approximating sequences are chosen in a careful manner (i.e. to be UASs), then accurate construction/reconstruction can be assured. We believe that wavelet constructions are likely to become a routine tool used in the simulation and solution of systems involving random components. Currently we are investigation implementation issues for the simulation of sample paths in variable step size ODE solvers using UASs and the correspondence principle.

References

[1] A. F. Karr, *Point Processes and Their Statistical Inference.* New York: Dekker, 1986.

[2] E. Masry, "Poisson sampling and spectral estimation of continuous-time processes," *IEEE Trans. on IT*, vol. 24, pp. 173–183, 1978.

[3] H. S. Shapiro and R. A. Silverman, "Alias-free sampling of random noise," *J. SIAM*, vol. 18, 1960.

[4] J. T. Gillis, *Multidimensional Point Processes and Random Sampling.* PhD thesis, UCLA, 1988.

[5] E. Masry, "Random sampling and reconstruction of spectra," *Information and Control*, vol. 19, pp. 275–288, 1971.

[6] E. Wong and B. Hajek, *Stochastic Processes in Engineering Systems.* Berlin: Springer-Verlag, 1985.

[7] N. Levinson, "Gap and density theorems," in *Am. Math. Soc. Colloq.*, vol. 26, Am. Math. Soc., 1940.

[8] J. J. Benedetto and W. Heller, "Irregular sampling and the theory of frames, I," *Note Math.*, vol. Köthe Memorial Volume, To appear, 1991.

[9] F. J. Beutler, "Error free recovery of signals from irregularly spaced samples," *SIAM Review*, vol. 8, pp. 99–105, 1966.

[10] A. Papoulis, *Probability, Random Variables and Stochastic Process, 2ed.* New York: McGraw–Hill, 1984.

[11] P. A. W. Lewis, *Stochastic Point Processes: Statistical Analysis, Theory, and Applications.* New York: Wiley, 1972.

[12] A. V. Balakrishnan, "A note on the sampling principle for continuous signals," *IRE Trans. on IT*, vol. 3, pp. 143–146, 1957.

[13] J. J. Benedetto, "Class lectures," *Univ. Maryland, Math. 767*, Fall, 1991.

[14] C. Houdré, "Private communication," *Univ. Maryland*, Fall, 1991.

[15] R. J. Duffin and A. C. Schaeffer, "A class of nonharmonic Fourier series," *Transaction of the American Mathematical Society*, vol. 72, pp. 341–366, 1952.

[16] I. Daubechies, "The wavelet transform, time-frequency localization and signal analysis," *IEEE Transactions on Information Theory*, vol. 36, pp. 961–1005, 1990.

[17] T. Kato, *Perturbation Theory for Linear Operators, Second Edition.* New York: Springer–Verlag, 1984.

[18] A. E. Taylor and D. C. Lay, *Introduction to Functional Analysis, Second Edition.* Malabar, Florida: Robert E. Krieger Publishing Co., 1986.

[19] H. A. Gindler and A. E. Taylor, "The minimum modulus of a linear operator and its use in spectral theory," *Studia Math.*, vol. 22, pp. 15—41, 1962.

[20] J. Zemanek, "Geometric interpertation of the essential minimum modulus," in *Invariant Subspaces and Other Topics*, vol. OT6, pp. 225–227, Brikhauser, 1981.

[21] S. Mallat, "Multiresolution approximation and wavelet orthogonal bases of l_2," *Transactions of the American Mathematical Society*, vol. 315, 1989.

[22] G. H. Golub and C. F. V. Loan, *Matrix Computations, Second Edition.* Baltimore Maryland: The Johns Hopkins University Press, 1989.

[23] M. Reed and B. Simon, *Methods pf Modern Mathematical Physics IV: Analysis of Operators.* New York: Academic Press, 1979.

[24] M. H. A. Davis, *Linear Estimation And Stochastic Estimation.* London: Halsted Press, 1979.

[25] A. V. Balakrishnan, *Communication Theory.* New York: McGraw-Hill, 1968.

[26] J. L. Doob, *Stochastic Processes.* New York: Wiley, 1958.