9

In Search of Multiparameter Markov Processes

Eugene Wong

Department of Electrical Engineering
and Computer Science
University of California, Berkeley, California

9.1 Introduction

One of the problems that John Thomas and I first tackled concerned the transition density of Markov processes. Through the Fokker-Planck equation we were able to show that the polynomial expansions first noted by Barrett and Lampard [1] could be interpreted in terms of the Sturm-Liouville problem to which the Fokker-Planck equation gives rise. My interest in Markov processess, kindled then, has never waned.

Over the years I have worked on a number of aspects of Markov processes, but one topic, above all, has continued to both challenge and frustrate me. This is the topic of multiparameter Markov processess, or Markovian random fields. Once again, I have returned to the topic, and thanks to an appointment as Miller Professor I am able to work on it full time this year. Though perhaps somewhat premature, I would like to pose the questions that have continually challenged me, my outline and approach, and present a few preliminary results. In the process it may also be interesting to briefly review the history of this topic.

9.2 Markovian Random Fields.

Let $\{X_t,\ t \in T\}$ be a family of random variables defined on a fixed probability space (Ω, F, P) and parameterized by elements of a set T. When T is an interval of the real line, $\{X_t,\ t \in T\}$ is simply an ordinary process. When T is a subset of a multidimensional space (say $\mathbb{R}^n$), then $\{X_t,\ t \in T\}$ is called a random field or multiparameter process, or a process with multidimensional time.

The way to extend the definition of "Markov process" was first suggested by Lévy [2]. A random field $\{X_t,\ t \in T\}$ is said to be Markovian if whenever ∂D is a smooth surface separating T into a bounded part D^- and a possibly unbounded part D^+, then $t \in D^+$ and $t' \in D^-$ imply that X_t and $X_{t'}$ are conditionally independent given the boundary data $\{X_s,\ s \in \partial D\}$. If we identify D^- as the "past", D^+ as the "future", and ∂D as the "present", then being Markovian means the conditional independence of "past" and "future" given the "present". This interpretation makes the Lévy definition consistent with the definition in the one-dimensional case.

My original interest in random fields was to find good models for images. From that point of view, Lévy's definition is attractive. It promises to delineate a class of processes that would represent both tractable and realistic models. Markovian independence gives one enough independence for computation to be possible, but not so much as to render the sample functions too ill behaved to be realistic.

9.3 Brownian Motion

For the one-parameter case, a Brownian motion can be defined as a Gaussian process $\{B_t,\ 0 \leq t < \infty\}$ with zero mean and a covariance function

$$EB_tB_s = \min(t,s). \tag{9.1}$$

The parameter space can be extended to the entire real line by modifying the covariance function to read

$$R(t,s) = EB_tB_s = (|t| + |s| - |t-s|)/2. \tag{9.2}$$

Lévy [2] defined a Brownian motion with $T = \mathbb{R}^n$ as a random field $\{B_t,\ t \in \mathbb{R}^n\}$ with zero mean and covariance function

$$R(t,s) = (|t| + |s| - |t-s|)/2 \tag{9.3}$$

where $|t|$ now denotes not the absolute value but the Euclidean norm

$$|t| = \left(\sum_i t_i^2\right)^{1/2}.$$

Lévy conjectured that so defined, the Brownian motion had some kind of Markovian property if the dimension of t was odd, but none if the dimension was even.

Lévy's conjecture was made precise and verified by McKean [3], who showed that for $n = 2k + 1$, B_t has the following Markovian property: Given the value of B_t and its $k+1$ "normal derivatives" on the boundary, ∂D, its "past"and "future" are indeed conditionally independent. Two points in McKean's proof are particularly important. First, the highest-order "normal derivative" needed on the boundary is always a generalized process so that the boundary data need to be defined with some care. Second, the source of the Markovian property for the Brownian motion with an odd-dimensional parameter appears to lie in the fact that for $n = 2k + 1$ the covariance function as given by (9.3) satisfies the equation

$$\triangle^k R(t,s) = A_k\delta(t-s) \tag{9.4}$$

where $\triangle$ is the Laplacian operator. Both the need to deal with generalized processes and the connection with the Laplacian operator are recurrent themes in Markovian random fields.

Gangoli [4] has considered generalizations of the Brownian motion for the cases where T is a Riemannian manifold and the covariance function is defined by a modified version of (9.3) where $|t-s|$ is replaced by the Riemannian distance $d(t,s)$. A natural question is: Does the resulting Brownian motion have any Markovian property? The answer is almost certainly yes if the parameter space has a Laplace-Beltrami operator, as for example in the case of spaces with constant curvature [5].

9.4 Ornstein-Uhlenbeck Processes

A one-parameter Ornstein-Uhlenbeck process (suitably normalized) can be defined as a Gaussian process with zero mean and a covariance function

$$R(t,s) = e^{-|t-s|}, \quad -\infty < t,s < \infty. \tag{9.5}$$

What would be a natural generalization for $T = \mathbb{R}^n$? The answer is provided as a part of the answer to a little different question.

Consider a zero-mean random field $\{X_t,\ t \in \mathbb{R}^n\}$ such that

$$R(t,s) = EX_tX_s = EX_{\tau\cdot t}X_{\tau\cdot s} \tag{9.6}$$

for all Euclidean motions τ (rotations and translations). It is easy to show that in this case $R(t,s)$ must be a function of just $|t-s|$. Furthermore, if X_t is quadratic mean continuous then $R(t,s)$ must have the form

$$R(t,s) = \int_0^\infty \frac{J_{(n-2)/2}(\lambda|t-s|)}{(\lambda|t-s|)^{(n-2)/2}} F(d\lambda) \tag{9.7}$$

where F is a finite Borel measure known as the spectral distribution of the process. In the late 1960s I posed the question: What must $F(\cdot)$ be in order that the process X_t be Markovian? It turned out that with a strict interpretation the question was not an interesting one. For $n \geq 2$ there is no finite, measure F for which the corresponding process is Markovian. However, if we relax the condition that F be finite then

$$F(d\lambda) = \frac{\lambda^{n-1}\, d\lambda}{\beta^2 + \lambda^2} \tag{9.8}$$

indeed yields a Markov process. However, (9.8) implies that

$$R(t,s) = A_n \frac{K_{(n-2)/2}(\beta\,|t-s|)}{(\beta\,|t-s|)^{(n-2)/2}} \tag{9.9}$$

and X_t must be considered a generalized process for $n \geq 2$. What does it mean then to say that X_t is Markovian? The answer is roughly as follows: Let D be a smooth $(n-1)$ surface. Although X_t is not well defined as a random variable at each point $t \in \partial D$, the surface integral of X_t on any subset A of ∂D is well defined as a second-order random variable $X(A)$. $\{X(A),\, A \in \partial D\}$ then represents the boundary data. To make this argument precise, we need to define X_t as a generalized process and define $X(A)$ by using an approximating sequence of testing functions [5].

In 1973, Nelson [6] independently proposed the Gaussian random field defined by (9.8) in the context of constructive quantum field theory, and called it the "Euclidean free field". One of the outstanding problems in a constructive quantum field theory is to construct non-Gaussian random fields that are isotropic and homogeneous (i.e., distributions are invariant under rotations and translations) and Markovian. To date, the success is limited.

9.5 Wiener Processes

In one dimension a Brownian motion or a Wiener process can be viewed as a zero-mean Gaussian process with a covariance function $\min(t,s)$ or as the indefinite integral of a Gaussian white noise. It is easy to define a Gaussian white noise η_t for $t \in \mathbb{R}^n$, and using it, we can define a Wiener process W by the integral

$$W_t = \int_{A_t} \eta_s \, ds \tag{9.10}$$

where A_t is the n-dimensional rectangle with the origin and t as two of its corners.

Of course, (9.10) is not properly an integral but merely a symbolic expression of the fact that W is more appropriately defined as a set-parametered process $W(A)$ which is independent on disjoint sets. To be specific let $W(A)$ be a Gaussian process parameterized by Borel sets A with $EW(A) = 0$ and $EW(A)W(B) =$ volume $(A \cap B)$. Now a Wiener process can be defined as $W(A_t)$.

For $t,s \in \mathbb{R}^n$ define the partial ordering $t > s$ by

$$t > s \Leftrightarrow t_i \geq s_i \text{ for all } i.$$

Then W_t is a martingale with respect to this partial ordering, and as such is the basis for a theory of stochastic integration that has been developed since 1974 [7,8]. Here we are less interested in the martingale property of W than in any Markovian property it may possess. Thus we would be interested in stochastic integration only if it has something to do with Markovian random fields. As we shall see later, such a connection indeed exists.

Meanwhile, a natural question is: Is W_t a Markov process? Surprisingly, the answer is "no". To see this, consider $\mathbb{R}^2$ and a triangular domain D^- bounded by the 45-degree line $\partial D = \{(\alpha, 1-\alpha), 0 \leq \alpha \leq 1\}$ as shown in Figure 9.1.

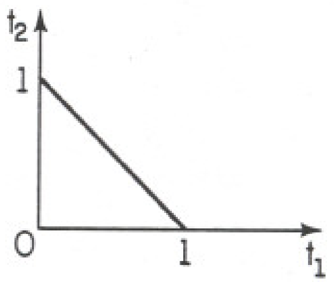

Figure 9.1

Now $W(1,1) = W(D^-) + W(A_{(1,1)} \cap D^+)$ and the second term is independent of $\{W_t, t \in D^-\}$. For W to be Markovian (in the Lévy sense) we need

$$E[W(D^-) \mid W_t,\ t \in \partial D] = W(D^-).$$

By a simple projection computation, we can show that

$$E[W(D^-) \mid W_t,\ t \in \partial D] = 2 \int_0^1 W_{u,u-1}\, du$$

which is definitely not equal to $W(D^-)$ [9].

It is indeed surprising that the Wiener process W_t, as the indefinite integral of white noise, is not Markovian. However, once we have discovered that it is not, it is not too difficult to conjecture as to the reason. Intuitively, "Markov" should be one derivative away from "white." In $\mathbb{R}^n$, an integral of white noise is n derivatives, not one, from it. Of course, this very vague intuitive notion needs clarification, but to do so needs mathematical machinery. To deal with differentiation in $\mathbb{R}^n$ requires a theory of differential forms, and in our case, stochastic differential forms. Put in another way: We need to be able to integrate on r-dimensional sets in $\mathbb{R}^n$ $(r \leq n)$ and not merely on n-dimensional sets. Thus the theory of stochastic integration associated with martingales that has been developed in recent years is not irrelevant, but inadequate.

9.6 In Search of a White Noise Connection

Consider an Ornstein-Uhlenbeck process on $\mathbb{R}^3$ with a covariance function given by (9.9). The case of $\beta = 0$ is an acceptable case, and it gives

$$R(t,s) = A \frac{1}{|t-s|}. \tag{9.11}$$

Now consider a vector $Z_t = (Z_t^1, Z_t^2, Z_t^3)$ of independent and identically distributed Ornstein-Uhlenbeck processes each with a covariance function

$$R(t,s) = \frac{1}{|t-s|}. \tag{9.12}$$

Let $\eta_t = \nabla Z_t$ be the divergence of Z_t (considered as a generalized process).

Then η_t is a Gaussian white noise. This provides a way of whitening an Ornstein-Uhlenbeck process that is not available without embedding it in a vector of independent processes [10]. We shall reexpress this relationship in terms of differential forms later, and use it to demonstrate explicitly the Markovian nature of an Ornstein-Uhlenbeck process.

The vector process Z_t is also related to Lévy's Brownian motion in a simple and geometrically suggestive way. Itô [11] has shown that a generalized random vector field such as Z_t can be uniquely decomposed into irrotational (curl = 0) and solenoidal (gradient = 0) components. Denote the irrotational component of Z_t by Z_{it}. Then

$$Z_{it} = \nabla B_t \tag{9.13}$$

where B_t is a Lévy's Brownian motion. Since the divergence of the solenoidal component is zero, we also have

$$\nabla \cdot \nabla B_t = \triangle B_t = \eta_t. \tag{9.14}$$

That is, the Laplacian of B_t is a white noise.

The connection among Gaussian white noise, the O-U process and Lévy's Brownian motion affirms our belief that Markovian random fields "come from" white noise fields and the connection is through geometric differentiation operations. It also points a possible way to the construction of non-Gaussian Markovian fields such as these needed in quantum field theory. But before such construction can be developed, we need a calculus of stochastic differential forms.

9.7 In Search of a Stochastic Calculus.

Let W_t be a one-parameter Wiener process, and consider a stochastic differential equation of the Itô type

$$dX_t = m(X_t,t)\,dt + \sigma(X_t,t)\,dW_t. \tag{9.15}$$

Under quite general conditions the solution X_t is a sample continuous Markov process, and this provides a way of constructing a large class of Markov processes in the one-dimensional case. To generalize the technique requires several things. First of all, it is not clear which process should play the role of W_t. There are at least three candidates: Lévy's Brownian motion B_t defined in Section 9.3, the Wiener process W_t defined in Section 9.5, and the Ornstein-Uhlenbeck process defined in Section 9.4 or possibly the vector version defined in Section 9.6. Second, it is not clear how the differential operator d is to be defined, and once defined how the term $\sigma\ dW$ is to be interpreted.

Intuitively, what I think we need is a definition for differential forms that would simultaneously deal with differentials of fields that are not strictly differentiable and with nonlinear operations that cannot be handled by the theory of generalized processes. This is exactly what the Itô calculus achieves in one dimension. We need its generalization to $\mathbb{R}^n$.

To give the basic ideas of what I think is needed, let us confine ourselves to the two-dimensional case $\mathbb{R}^2$. Define oriented r-rectangles $(r = 1,2)$ as follows: A 1-rectangle is a line segment parallel to one of the two axes, and each 1-rectangle has one of two orientations. A 2-rectangle is just an ordinary rectangle with sides parallel to the two axes and has two possible orientations (pointing out and pointing in, say). Finally, a 0-rectangle is just a point and it too is given two possible orientations. If A is an oriented rectangle, then $-A$ is the same rectangle with the opposite orientation.

For a number of reasons we need to consider linear combinations of the form

$$\sum_k \alpha_k A_k$$

where α_k are real numbers and A_k are oriented rectangles of the same dimension. The linear combinations satisfy certain natural axioms [12] and are known as r-chains ($r = 0,1,2$). Observe that the boundary of an oriented r-rectangle is a sum of oriented $(r-1)$-rectangles. Hence the boundary of an r-chain is an $(r-1)$-chain.

A two-dimensional set in $\mathbb{R}^2$ can be approximated by a sequence of 2-chains (by subdivision, for example). A one-dimensional curve in $\mathbb{R}^2$ can also be approximated by a sequence of 1-chains (by staircase-like approximations, for example). Now suppose that X is a linear map of the space of r-chains into a space of square-integrable random variables. With appropriate continuity conditions, X can be extended to all sets that can be approximated by r-chains. Intuitively, $X(A)$ for an r-dimensional set A can be thought of as an integral

$$X(A) = \int_A \xi$$

where ξ is a random differential r-form, except that for many interesting cases ξ is only a generalized r-form (r-current). We shall call X a random r-cochain (or equivalently a random r-form).

The exterior differential of a random r-cochain is an $(r+1)$-cochain defined as follows:

$$(dX)(A) = X\ (\partial A) \tag{9.16}$$

when ∂A is the boundary of A. This is nothing but the Stoke's theorem, used here as a definition rather than as a property. Now take a Wiener process W_t. It can be considered as a 0-form. Its exterior derivative dW is a 1-form, so that dW is parameterized by one-dimensional curves. If we define d_1W and d_2W by

$$\begin{aligned} d_1W &= dW && \text{on horizontal line segments} \\ &= 0 && \text{on vertical line segments} \end{aligned}$$

and

$$\begin{aligned} d_2W &= dW && \text{on vertical line segments} \\ &= 0 && \text{on horizontal line segments} \end{aligned} \tag{9.17}$$

then

$$dW = d_1W + d_2W$$

and d_iW are both 1-forms. It is easy to show that

$$d(d_2W) = -d(d_1W) = \eta$$

is a white noise 2-form, and because of it both d_1W and d_2W are Markovian. We now see that although W_t is not Markovian, its exterior derivative dW is Markovian in the sense that its horizontal and vertical components d_iW are both Markovian.

Thus far everything is linear. To proceed further, we need to introduce the exterior product $X \wedge Y$ (between r and p forms) which is a nonlinear operation. For example, $d_1W \wedge d_2W$ is a 2-form that takes the value

$$W(dt_1,t_2)\ W(t_1,dt_2)$$

on an incremental rectangle $dt_1 \wedge dt_2$ at (t_1,t_2). It turns out that the exterior product is closely related to the type-2 stochastic integral introduced in [7] and the multiple Itô integral introduced in [13].

In the theory of stochastic integration developed for two parameter martingales, one source of difficulty has been the lack of a useful calculus. For example, the differentiation formula (expressed as a transformation of integrals into integrals) derived in [14] is difficult to use. With the introduction of stochastic differential forms, a much simpler calculus is beginning to emerge. For example, let W be a two-parameter Wiener process. Let f be a twice continuously differentiable function. Then $f(W)$ is a 0-form and its exterior derivative is a 1-form given by

$$df(W_t) = f'(W_t) \wedge dW_t + 1/2 f''(W_t) \wedge d\nu_t \tag{9.18}$$

where ν_t is just the area of A_t (the rectangle bounded by the origin and t), d denotes exterior differentiation and $\wedge$ denotes exterior product.

The introduction of differential forms also makes possible certain transformations that should preserve Markovian properties in a way analogous to the Itô differential equations. As an example, consider an equation

$$dX_t = X_t \wedge dt + dW_t. \tag{9.19}$$

As a global equation, it has no solution. We can see this by noting that $ddX_t = 0$ (dd of any form is 0) but $d(X_t \wedge dt + dW_t) = dX_t \wedge dt \neq 0$.

However, the equation has a solution on any path if we require it to be satisfied only on that path. Now, suppose that we take a collection of paths $\Gamma = \{\gamma\}$ such that no two paths in Γ ever cross and collectively the paths cover the entire space $\mathbb{R}^2$. Then solving the differential equation on each path yields a 0-form X_t on $\mathbb{R}^2$. For example, let Γ be the set of radial paths, i.e.,

$$\gamma_\theta = \{(r\cos\theta, r\sin\theta),\ \ 0 \leq r \leq \infty\},\ \ 0 \leq \theta \leq 2\pi$$

then, on γ_θ (9.19) becomes

$$X(dr,\theta) = X(r,\theta)dr + W(dr\cos\theta, dr\ \sin\theta)$$

and the solution is

$$X(r,\theta) = e^r X_0 + \int_0^r e^{-(r-u)} W(du\ \cos\theta,\ dr\ \sin\theta).$$

This solution can be written in a coordinate-free form as

$$X_t = e(t)\ [\ X_0 + \int_{\gamma_t} e^{-1}\ dW\ ]$$

where ρ is a 0-form defined by $e(t) = \exp|t|$ and γ_t is the radial path from the origin to t.

I believe that equations such as (9.19), solved locally on paths, provide a way of transforming processes that preserves Markovian properties. The extent to which this is true is under active investigation.

9.8 Conclusion

The development of a theory of Markovian random fields faces a number of obstacles. On the one hand, in any dimension higher than 1 the Markovian property appears to be incompatible with sample continuity, so that one has to deal with generalized processes. At the same time, the very nature of the Markovian property is local, so that generalized processes per se are not a suitable model. Furthermore, to construct non-Gaussian Markov processes requires nonlinear operations that cannot be handled by the existing theory of generalized processes. I am convinced that what is needed is a stochastic calculus like the one Itô developed for one dimension, but one that is necessarily differentio-geometric in higher dimensions.

References

[1] J.F. Barrett and D.G Lampard, An expansion for some second-order probability distributions and its applications to noise problems, *IRE Trans. on Inform. Theory*, Vol. IT-1, pp. 10-15, 1955.

[2] P. Lévy, A special problem of Brownian motion, and a general theory of Gaussian random functions, *Proc. 3rd Berkeley Symp. Math. Stat. and Prob.*, Vol. 2, Berkeley, CA: Univ. of California Press, pp. 133-175. 1956.

[3] H.P. McKean Jr., Brownian motion with a second dimensional term, *Theory Prob. Appl.*, Vol. 8, pp. 335-365, 1963.

[4] R. Gangoli, Abstract harmonic analysis and Lévy's Brownian motion of several parameters, *Proc. 5th Berkeley Symp. Math. Stat. and Prob., Vol. II-1*, Berkeley, CA: Univ. of California Press, Vol. II-1, pp. 13-30, 1967.

[5] E. Wong, Homogeneous Gauss-Markov random fields, *Ann. Math. Statist.*, Vol. 40, pp. 1625-1634, 1969.

[6] E. Nelson, The free Markov field, *J. Functional Analysis*, Vol. 12, pp. 211-227, 1973.

[7] E. Wong and M. Zakai, Martingales and stochastic integrals for processes with a multidimensional parameter, *Z. Wahr. verw. Geb.*, Vol. 29, pp. 109-122, 1974.

[8] R. Cairoli and J. Walsh, Stochastic integrals in the plane, *Acta Math.* Vol. 134, pp. 111-183, 1975.

[9] J. Walsh, Martingales with a multidimensional parameter and stochastic integrals in the plane, *Notes for cours de troisiéme cycle, University of Paris, VI*, 1977.

[10] C. Belisle, Private communication, 1983.

[11] K. Itô Isotropic random current, *Proc. 3rd Berkeley Symp. Math. Stat. and Prob., Vol. 2*, Berkeley, CA: Univ. of California Press, pp. 125-132, 1956.

[12] H. Whitney, *Geometric Integration Theory,* Princeton, NJ: Princeton University Press, 1957.

[13] B. Hajek and E. Wong, Set parametered martingales and multiple stochastic integration, *Lecture Notes in Mathematics, No. 851,* Berlin: Springer-Verlag, pp. 119-151, 1980.

[14] E. Wong and M. Zakai, Differentiation formulas for stochastic integrals in the plane, *Stoch. Proc. Appl.*, Vol. 6, pp. 339-349, 1978.