Mathematica Slovaca

Jaroslav Mohapl

On estimation in random fields generated by linear stochastic partial differential equations

Mathematica Slovaca, Vol. 49 (1999), No. 1, 95--115

Persistent URL: `http://dml.cz/dmlcz/131241`

Math. Slovaca, **49** (1999), No. 1, 95–115

Mathematica Slovaca

ON ESTIMATION IN RANDOM FIELDS GENERATED BY LINEAR STOCHASTIC PARTIAL DIFFERENTIAL EQUATIONS

Jaroslav Mohapl

(*Communicated by Milan Medveď*)

ABSTRACT. A stationary random field with a rational spectral density function is often associated with a stochastic partial differential equation (SPDE). The question motivating this study is whether and how knowledge of the SPDE may simplify the statistical analysis of the associated random field.

1. Introduction

A linear stochastic partial differential equation (SPDE) is described by the relation

$$\mathcal{P}_\theta(\partial)c = \sigma\varepsilon\,, \tag{1}$$

where $\mathcal{P}_\theta(\partial)$ is a formal linear partial differential operator, c and ε are generalized random fields, or equivalently, distribution-valued processes and θ, σ are unknown parameters. In general, the problem is to estimate θ and σ and to find criteria for goodness of fit between the model and observed data. This article explains the mathematical meaning of (1), the relation between the solution of (1) and the random field models used in statistics and how the solution of (1) is applicable for the analysis of spatial data, in particular for parameter estimation.

A classical example motivating study of SPDE's is in [17] (Whittle, 1962). In the analysis of a wheat yield data set, Whittle used the equation

$$\partial_t\xi = \frac{1}{2}(\partial^2_{x_1}\xi + \partial^2_{x_2}\xi) - \theta\xi + \sigma Z \tag{2}$$

with Z interpreted as a zero-mean white noise. The solution ξ of (2) was claimed to be a random field with spatial covariance function described by a modified

AMS Subject Classification (1991): Primary 62M30; Secondary 60G60.
Key words: maximum likelihood, spatial process, generalized random field, Gauss process, Schwartz distribution.

Bessel function of the second kind, order zero:

$$R_{\theta,\sigma}(x) = \frac{\sigma^2}{2\pi} K_0\big(|x|\sqrt{2\theta}\big)\,, \tag{3}$$

where $x \in \mathbb{R}^2$ and $|\cdot|$ is the Euclidean norm of x. We recall that $R_{\theta,\sigma}(x-y) = E_{\theta,\sigma}\xi(t,x)\xi^*(t,y)$ for each $t>0$ and pair $x,y \in \mathbb{R}^2$. The covariance is derived from (2) heuristically by means of the Fourier transform method. Whittle assumed that the wheat yield depends on the fertility of the soil. The random variable $\xi = \xi(t,x,\omega)$ denotes the amount of nutrients in a unit of soil at time t and location $x = (x_1, x_2)$. The event ω from a probability field is usually omitted in the notation. Equation (2) says that the change in the amount of nutrients during a time increment on the left is proportional to the change of the concentration gradient along the x_1 and x_2 axes minus a specific discharge due to the consumption of nutrients by the wheat plus a random term describing the unequal distribution of the nutrients caused by heterogeneity of the soil structure. The magnitude θ of the discharge as well as variability of the soil inhomogeneities determined by σ must be estimated. Many other applications using SPDE's may be found in [13] (Namachchivaya, 1988).

The heuristic approach used in Whittle's paper has at least two disadvantages. A random field with covariance function (3) has an infinite variance, because $K_0(x)$ behaves as $-\ln(|x|)$ for small x. Consequently, ξ cannot be a Gaussian process. Second, the trajectories of ξ are not differentiable and therefore do not satisfy (2). Hence, compared to ordinary autoregressions, equation (2) does not provide any residuals and as a tool for study of the process is useless. The next section outlines a way how to by-pass the differentiability problem.

2. The general linear SPDE

A linear SPDE obtained from physical considerations has form

$$\sum_{|k|\le p} a_k(\theta)\partial_k \xi(x)\,\mathrm{d}x = \sigma Z(\mathrm{d}x)\,, \tag{4}$$

where $x \in (-\infty,\infty)^d$, $\partial_k = \partial^{|k|}/\partial_{x_1}^{k_1}\dots\partial_{x_d}^{k_d}$ are mixed partial derivatives, $k_1,\dots,k_d$ and p are non-negative whole numbers, $|k| = k_1 + \dots + k_d \le p$, $k = (k_1,\dots,k_d)$, ξ is a random field and Z is an orthogonal random measure. The volume $\mathrm{d}x = \mathrm{d}x_1\dots\mathrm{d}x_d$ emphasizes that the right hand side of (4) is the density of an orthogonal random measure with respect to the ordinary Euclidean measure living on subsets of $(-\infty,\infty)^d$. If we consider the derivatives as variables of a polynomial $\mathcal{P}_\theta(\partial)$ then we may write (4) in the equivalent form

$$\mathcal{P}_\theta(\partial)\xi(x)\,\mathrm{d}x = \sigma Z(\mathrm{d}x)\,. \tag{5}$$

The polynomial $\mathcal{P}_\theta(\partial)$ is called a formal linear partial differential operator with constant coefficients. One reason for using the word formal is that the trajectories of ξ are usually not integrable. The functional (or distribution) valued equation (1) is obtained from (5) as follows. We multiply both sides of (5) by a "test" function ϕ which is sufficiently smooth and integrable. We integrate over $\mathbb{R}^d = (-\infty, \infty)^d$ (by parts on the left side) and obtain the equation

$$\int_{\mathbb{R}^d} \xi(x)\mathcal{P}'_\theta(\partial)\phi(x)\, \mathrm{d}x = \sigma \int_{\mathbb{R}^d} \phi(x)\, Z(\mathrm{d}x)\,, \tag{6}$$

where $\mathcal{P}'_\theta(\partial)$ is called the formal adjoint operator to $\mathcal{P}_\theta(\partial)$. The operator $\mathcal{P}'_\theta(\partial)$ is obtained from $\mathcal{P}_\theta(\partial)$ by integration by parts.

Definition 1. Let $\mathcal{S}$ be the set of infinitely differentiable functions with compact support in $\mathbb{R}^d$. A process ξ that satisfies (6) for all $\phi \in \mathcal{S}$ almost surely is called a *solution* to equation (5).

Definition 1 removes any differentiability assumptions on ξ. For mathematical operations it is convenient to introduce random elements c and ε defined by the relations $c(\phi) = \int_{\mathbb{R}^d} \xi(x)\phi(x)\, \mathrm{d}x$ and $\varepsilon(\phi) = \int_{\mathbb{R}^d} \phi(x)\, Z(\mathrm{d}x)$, respectively, and to write (6) in the form

$$c\big(\mathcal{P}'_\theta(\partial)\phi\big) = \sigma\varepsilon(\phi)\,. \tag{7}$$

If c and ε admit a modification to a generalized random field then (7) serves as the definition of (1). For more details see [5] (Itô, 1984) or [16] (Walsh, 1986).

Equation (6) is often considered in the more general form

$$\sum_{|k|\le p} a_k(\theta)\partial_k\xi(x)\,\mathrm{d}x = \sigma \sum_{|k|\le q} b_k(\theta)\partial_k Z(\mathrm{d}x)\,, \tag{8}$$

where $\sum_{|k|\le q} b_k(\theta)\partial_k = \mathcal{Q}_\theta(\partial)$ is also a formal linear partial differential operator with constant coefficients. In this case we can identify (8) with the relation

$$\int_{\mathbb{R}^d} \xi(x)\mathcal{P}'_\theta(\partial)\phi(x)\, \mathrm{d}x = \sigma \int_{\mathbb{R}^d} \mathcal{Q}'_\theta(\partial)\phi(x)\, Z(\mathrm{d}x) \tag{9}$$

and write briefly

$$\mathcal{P}_\theta(\partial)\xi(x)\,\mathrm{d}x = \sigma\mathcal{Q}_\theta(\partial)Z(\mathrm{d}x)\,. \tag{10}$$

If the left and right hand side of (10) admit modification to a generalized random field we may again describe (10) by (1), where we set, say, $\varepsilon = \mathcal{Q}_\theta(\partial)W$. We can use a suitable class $\mathcal{S}$ of test functions to extend Definition 1 of a solution to (9) accordingly.

A solution of (9) is usually assumed to be an ordinary stochastic process with the rational spectral density function

$$f_{\theta,\sigma}(\lambda) = \sigma^2 \frac{|\mathcal{Q}_\theta(\mathrm{i}\,\lambda)|^2}{|\mathcal{P}_\theta(\mathrm{i}\,\lambda)|^2}\,. \tag{11}$$

It is natural to consider such a process by analogy with discrete *ARMA* processes. The following theorem describes a class of solutions to the equation (8).

THEOREM 1. *If*

$$\int_{\mathbb{R}^d}\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} f_{\theta,\sigma}(\lambda)\,\mathrm{e}^{\mathrm{i}\langle x-y,\lambda\rangle}\,\mathrm{d}\lambda\,\mathrm{d}x\,\mathrm{d}y < \infty \tag{12}$$

and Z *is an orthogonal measure with* $E\big(Z(\mathrm{d}x)\big)^2 = \mathrm{d}x$ *then the process*

$$\xi(x) = \sigma \int_{\mathbb{R}^d} \mathrm{e}^{\mathrm{i}\langle x,\lambda\rangle} \int_{\mathbb{R}^d} \frac{\mathcal{Q}_\theta(\mathrm{i}\,\lambda)}{\mathcal{P}_\theta(\mathrm{i}\,\lambda)}\,\mathrm{e}^{-\,\mathrm{i}\langle y,\lambda\rangle}\; Z(\mathrm{d}y)\,\mathrm{d}\lambda \tag{13}$$

satisfies (9) *for every* ϕ *with compact support in* $\mathbb{R}^d$ *almost surely.*

Proof. Using (12) and the well-known Plancherel theorem proved e.g. in [19] (Yosida, 1974) we can verify that ξ is a well defined stochastic process with finite variance and spectral density (11). Trajectories of the process may be considered integrable. If we substitute (13) into the right hand side of (9) and use the stochastic version of Fubini's theorem in [7] (Liptser and Shiryayev, 1977) then we arrive to the stochastic integral $\sigma \int_{\mathbb{R}^d} g(y)\,Z(\mathrm{d}y)$, where $g(y)$ denotes the expression

$$\int_{\mathbb{R}^d} \mathrm{e}^{-\,\mathrm{i}\langle y,\lambda\rangle}\,\frac{\mathcal{Q}_\theta(\mathrm{i}\,\lambda)}{\mathcal{P}_\theta(\mathrm{i}\,\lambda)} \int_{\mathbb{R}^d} \mathrm{e}^{\mathrm{i}\langle x,\lambda\rangle}\,\mathcal{P}'_\theta(\partial)\phi(x)\;\mathrm{d}x\,\mathrm{d}\lambda$$

$$= \int_{\mathbb{R}^d} \mathrm{e}^{-\,\mathrm{i}\langle y,\lambda\rangle}\,\mathcal{Q}_\theta(\mathrm{i}\,\lambda) \int_{\mathbb{R}^d} \mathrm{e}^{\mathrm{i}\langle x,\lambda\rangle}\,\phi(x)\;\mathrm{d}x\,\mathrm{d}\lambda = \mathcal{Q}'_\theta(\partial)\phi(y)\,. \tag{14}$$

Thus (9) is satisfied. □

The commonly used process

$$\xi(x) = \sigma \int_{\mathbb{R}^d} \frac{\mathcal{Q}_\theta(\mathrm{i}\,\lambda)}{\mathcal{P}_\theta(\mathrm{i}\,\lambda)}\,\mathrm{e}^{\langle x,\lambda\rangle}\; Z(\mathrm{d}\lambda) \tag{15}$$

does not satisfy (9). However, it is easy to verify that if Z is a Gaussian white noise orthogonal measure, i.e. $Z(\mathrm{d}x) \sim \mathcal{N}(0,\mathrm{d}x)$, and the trajectories of (15) are integrable over compact subsets of $\mathbb{R}^d$ then, for a given ϕ, the random variable obtained by substitution of (15) into the left hand side of (9) has the same probability distribution as the random variable on the right. This can be utilized for the analysis of residuals. It is also important because processes used in statistics rather rarely satisfy condition (12).

3. Schwartz distributions

The class of SPDE's with solution ξ in the sense of Definition 1 is fairly narrow. But if we give up the requirement that $c(\phi) = \int_{\mathbb{R}^d} \phi(x)\xi(x)\,\mathrm{d}x$ and look only for a random linear functional c that satisfies (7) (i.e. (1)) for all $\phi \in \mathcal{S}$ then we can derive the following.

Theorem 2. *Let $\mathcal{S}$ consist of rapidly decreasing functions with Schwartz topology and let $\mathcal{S}'$ be the topological dual of $\mathcal{S}$. Then for each linear partial differential operator $\mathcal{P}_\theta(\partial)$ with constant coefficients and for each $\mathcal{S}'$-valued random element ε, there exists an $\mathcal{S}'$-valued random element c such that $\mathcal{P}_\theta(\partial)c = \sigma\varepsilon$ almost surely.*

Proof. If $\mathcal{P}_\theta(\partial)$ is a linear partial differential operator with constant coefficients then, according to the Malgrange-Ehrenpreis Theorem in [19] (Yosida, 1974), for each value of ε in $\mathcal{S}'$ there is an element $c \in \mathcal{S}'$ such that $\mathcal{P}_\theta(\partial)c = \varepsilon$. Hence, $\mathcal{P}_\theta(\partial)$ is a one to one mapping from $\mathcal{S}'$ onto $\mathcal{S}'$ with inverse $\mathcal{P}_\theta^{-1}(\partial)$. The operator $\mathcal{P}_\theta(\partial)$ is continuous in the topology of $\mathcal{S}'$ and therefore, it is Borel measurable. By [14; Corollary 24.25] (Parthasarathy, 1978), there exists a set of probability one in $\mathcal{S}'$ such that $\mathcal{P}_\theta(\partial)$ restricted to this set becomes bimeasurable and the relation $c = \mathcal{P}_\theta^{-1}(\partial)\varepsilon$ defines the desired random element. □

Notice that Theorem 2 assumes only that ε has values in $\mathcal{S}'$ almost surely but there are no restrictions on the probability distribution of the process.

Definition 2. We call the element c in Theorem 2 the *distribution-valued solution* of (1).

In this context, distribution means a Schwartz distribution. For its definition and properties see e.g. [19] (Yosida, 1974). In statistical literature, c is more usually referred to as a generalized random field. The part played by the process c is comparable to that of complex numbers in algebra. Only a limited number of algebraic equations have roots in the real domain but each of them has complex roots. These, however, may not be observable if only real-valued measurements are sampled. Similarly, at this stage, values of c may be reconstructed from ξ only if the representation $c(\phi) = \int_{\mathbb{R}^d} \phi(x)\xi(x)\,\mathrm{d}x$ is valid for each test function ϕ.

By definition, a distribution valued random element W is a white noise if $EW(\phi)W(\psi) = \int_{\mathbb{R}^d} \phi(x)\psi(x)\,\mathrm{d}x$ for each pair $\phi, \psi \in \mathcal{S}$. The element W is called Gaussian if the vector $\big(W(\phi_1), \ldots, W(\phi_N)\big)$ has a normal distribution for each N-tuple $\phi_1, \ldots, \phi_N \in \mathcal{S}$. The next proposition is proved in several papers and monographs. See for example [5] (Itô, 1984) or [16] (Walsh, 1986).

PROPOSITION 1. *Let $\mathcal{S}$ consist of the rapidly decreasing functions with Schwartz topology. Each stochastic orthogonal measure Z such that $EZ(\mathrm{d}x) = 0$ and $E\big(Z(\mathrm{d}x)\big)^2 = \mathrm{d}x$ defines, by the relation $\phi \mapsto \int \phi(x)\, Z(\mathrm{d}x)$, an $\mathcal{S}'$ valued random element $W(\phi)$ with the properties $EW(\phi) = 0$ and $EW(\phi)W(\psi) = \int \phi(x)\psi(x)\, \mathrm{d}x$ for each pair ϕ and ψ in $\mathcal{S}$.*

COROLLARY 1. *If Z is the orthogonal measure from Proposition 1 then equations (4) and (8) may be considered as equalities between distributions and they have at least one distribution-valued solution c.*

Corollary 1 is a direct consequence of Theorem 2 and Proposition 1. It also justifies the effort we spent in the previous section by describing equations (4) and (8) using differential operators an linear forms defined by integrals. See equations (1) and (10).

The relation $\phi \mapsto E_{\theta,\sigma}c(\phi)$ defines a distribution called the mean of c. If there is a continuous linear operator $\Gamma_{\theta,\sigma}$ from $\mathcal{S}$ into $\mathcal{S}'$ such that $(\Gamma_{\theta,\sigma}\phi)(\psi) = E_{\theta,\sigma}\big(c(\phi) - E_{\theta,\sigma}c(\phi)\big)\big(c(\psi) - E_{\theta,\sigma}c(\psi)\big)$ for each possible ϕ and ψ in $\mathcal{S}$, then $\Gamma_{\theta,\sigma}$ is called the covariance operator of c. If the operator $\mathcal{P}_\theta(\partial)$ has an inverse $\mathcal{P}_\theta^{-1}(\partial)$ for each θ and W is a Gaussian white noise then c is also Gaussian with covariance

$$\Gamma_{\theta,\sigma} = \sigma^2\big(\mathcal{P}_\theta'(\partial)\mathcal{P}_\theta(\partial)\big)^{-1}, \tag{16}$$

where -1 denotes the inverse operator. The covariance structure of c obviously resembles the structure we would obtain in the case of an ordinary autoregression model. More generally, if c is the distribution-valued process determined by the equation (10) then its covariance operator is

$$\Gamma_{\theta,\sigma} = \sigma^2\mathcal{P}_\theta^{-1}(\partial)\mathcal{Q}_\theta(\partial)\mathcal{Q}_\theta'(\partial)\mathcal{P}_\theta'^{-1}(\partial). \tag{17}$$

DEFINITION 3. Let $\Gamma_{\theta,\sigma}$ be a covariance operator of a generalized random field c. If there is a symmetric non-negative definite function $R_{\theta,\sigma}$ such that

$$\Gamma_{\theta,\sigma}\phi(x) = \int_{\mathbb{R}^d} R_{\theta,\sigma}(x,y)\phi(y)\, \mathrm{d}y \tag{18}$$

and

$$(\Gamma_{\theta,\sigma}\phi)(\psi) = \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} R_{\theta,\sigma}(x,y)\phi(y)\psi(x)\, \mathrm{d}y \tag{19}$$

for all test functions ϕ and ψ then $R_{\theta,\sigma}$ is called the *covariance function* of c.

If c is a solution of (1) then existence of the covariance function is obviously linked to existence of a process ξ which represents c in the sense that $c(\phi) = \int \xi(x)\phi(x)\, \mathrm{d}x$ for every test function ϕ.

THEOREM 3. *Let Z be a Gaussian orthogonal random measure with $E\big(Z(\mathrm{d}x)\big)^2 = \mathrm{d}x$ and let $R_{\theta,\sigma}(x)$ be a covariance function with rational spectral density* (11). *Denote by c the distribution-valued solution of* (8). *Then we can find a distribution-valued Gaussian process $\tilde{c}$ and a stationary Gaussian process ξ with spectral density* (11) *such that $\tilde{c}$ has the same probability distribution as c and for every test function $\tilde{c}(\phi) = \int_{\mathbb{R}^d} \xi(x)\phi(x)\,\mathrm{d}x$.*

Proof. The process c exists by Theorem 1. Let us define ξ by (15), which guarantees that ξ has covariance function $R_{\theta,\sigma}(x)$. As in the case of ordinary Gaussian processes, the theorem will be proved if we show that c and $\tilde{c}$ have the same covariance operator. For each pair ϕ, ψ of test functions

$$\begin{aligned}\int_{\mathbb{R}^d} \Gamma_{\theta,\sigma}\mathcal{P}'_\theta(\partial)\phi(x)\mathcal{P}'_\theta(\partial)\psi(x)\,\mathrm{d}x &= E_{\theta,\sigma}c\big(\mathcal{P}'_\theta(\partial)\phi\big)c\big(\mathcal{P}'_\theta(\partial)\psi\big)\\ &= E_{\theta,\sigma}W\big(\mathcal{P}'_\theta(\partial)\phi\big)W\big(\mathcal{P}'_\theta(\partial)\psi\big) \qquad (20)\\ = \sigma^2\int_{\mathbb{R}^d} \mathcal{Q}'_\theta(\partial)\,\phi(x)\mathcal{Q}'_\theta(\partial)\,\psi(x)\,\mathrm{d}x &= E_{\theta,\sigma}\tilde{c}\big(\mathcal{P}'_\theta(\partial)\phi\big)\tilde{c}\big(\mathcal{P}'_\theta(\partial)\psi\big)\,,\end{aligned}$$

where W is the $\mathcal{S}'$-valued white noise. The last equality follows from the Fourier transform representation of $R_{\theta,\sigma}$ and proves the assertion. □

Theorem 3 says that every covariance function on $\mathbb{R}^d$ with a rational spectral density can be associated with a distribution-valued solution of a SPDE with a Gaussian white noise right hand side. As a direct consequence of Theorem 3 we have:

COROLLARY 2. *Let $R_{\theta,\sigma}$ be a covariance function of a stationary random field with spectral density $f_{\theta,\sigma}(\lambda) = \sigma^2/|\mathcal{P}_\theta(\lambda)|^2$, where $\mathcal{P}_\theta$ is a polynomial whose coefficients are functions of θ. Then for every θ and σ our $R_{\theta,\sigma}$ satisfies the equation*

$$\mathcal{P}'_\theta(\partial)\mathcal{P}_\theta(\partial)R_{\theta,\sigma} = \sigma^2\delta\,, \qquad (21)$$

where δ is the Dirac distribution.

In the case of a so called evolution equation one can obtain a more detailed result:

PROPOSITION 2. *Consider the distribution-valued solution of the equation*

$$\big(\partial_t - \mathcal{A}_\theta(\partial_x)\big)\xi(t,x)\,\mathrm{d}t\,\mathrm{d}x = \sigma\mathcal{B}_\theta(\partial_x)Z(\mathrm{d}t,\mathrm{d}x)\,, \qquad (22)$$

where Z *is an orthogonal measure with* $E\big(Z(\mathrm{d}t,\mathrm{d}x)\big)^2 = \mathrm{d}t\,\mathrm{d}x$ *and* $\mathcal{A}_\theta(\partial_x)$ *and* $\mathcal{B}_\theta(\partial_x)$ *are linear differential operators acting only on the variable* x, $x \in \mathbb{R}^d$. *Let* $\mathcal{A}_\theta$ *have the property:*

i) *there is a negative constant* M *such that* $\mathcal{A}_\theta(\mathrm{i}\,\lambda) \le M$ *for all* $\lambda \in \mathbb{R}^d$ *and* θ *in the parameter set,*

and let

ii) $-|\mathcal{B}_\theta(\partial_x)|^2/\mathcal{A}_\theta(\mathrm{i}\,\lambda)$ *be a valid spectral function.*

Then the covariance function of c *exists and for* $t > 0$ *satisfies the equation*

$$\big(\partial_t - \mathcal{A}_\theta(\partial_x)\big)R_{\theta,\sigma}(t,x) = 0\,, \qquad R_{\theta,\sigma}(0,x) = \rho_{\theta,\sigma}(x) \tag{23}$$

with the initial condition

$$\rho_{\theta,\sigma}(x) = -\frac{\sigma^2}{(2\pi)^d}\int_{\mathbb{R}^d} \mathrm{e}^{\mathrm{i}\langle x,\lambda\rangle}\,\frac{|\mathcal{B}_\theta(\mathrm{i}\,\lambda)|^2}{\mathcal{A}_\theta(\mathrm{i}\,\lambda)}\,\mathrm{d}\lambda\,. \tag{24}$$

Proof. The fundamental solution $T_{t,x}$ of the equation (22) satisfies the relation $\mathcal{P}_\theta(\partial)T_{t,x} = \delta_t \otimes \delta_x$ with the initial condition $T_{0,x} = \delta_x$, where the symbol $\delta_t \otimes \delta_x$ denotes the direct product of two Dirac distributions acting on the variables t and x, respectively, and

$$\mathcal{P}_\theta(\partial) = \partial_t - \mathcal{A}_\theta(\partial_x)\,. \tag{25}$$

It is a distribution defined for $t > 0$ by means of the function

$$T(t,x) = H(t)(2\pi)^{-d}\int_{\mathbb{R}^d} \mathrm{e}^{\mathrm{i}\langle x,\lambda\rangle + t\mathcal{A}_\theta(\mathrm{i}\,\lambda)}\,\mathrm{d}\lambda\,, \tag{26}$$

where $H(t)$ denotes the Heaviside function. The integral converges for every $t > 0$ because

$$\left|\int_{\mathbb{R}^d} \mathrm{e}^{\mathrm{i}\langle x,\lambda\rangle + t\mathcal{A}_\theta(\mathrm{i}\,\lambda)}\,\mathrm{d}\lambda\right| \le \int_{\mathbb{R}^d} \mathrm{e}^{t\mathcal{A}_\theta(\mathrm{i}\,\lambda)}\,\mathrm{d}\lambda \le -\int_{\mathbb{R}^d} \frac{1}{t\mathcal{A}_\theta(\mathrm{i}\,\lambda)}\,\mathrm{d}\lambda < \infty \tag{27}$$

according to the assumption ii). If the covariance function of the stationary solution exists then it satisfies the equation

$$\Gamma_{\theta,\sigma}\phi(t,x) = \int_{\mathbb{R}^d} R_{\theta,\sigma}(t-s,x-y)\phi(s,y)\,\mathrm{d}s\,\mathrm{d}y \tag{28}$$

for every rapidly decreasing function $\phi \in \mathcal{S}$. Therefore

$$\mathcal{P}_\theta(\partial)\Gamma_{\theta,\sigma}\phi(t,x) = \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} \mathcal{P}_\theta(\partial)R_{\theta,\sigma}(t-s,x-y)\phi(s,y)\,\mathrm{d}s\,\mathrm{d}y \tag{29}$$

and since $\Gamma_{\theta,\sigma} = \left(\mathcal{P}'_\theta(\partial)\mathcal{P}_\theta(\partial)\right)^{-1}$, we have $\mathcal{P}_\theta(\partial)\Gamma_{\theta,\sigma} = \mathcal{P}'^{-1}_\theta(\partial)$. The operator $\mathcal{P}'^{-1}_\theta(\partial)$ is defined by the relation

$$\mathcal{P}'^{-1}_\theta(\partial)\phi(t,x) = \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} T'(t-s, x-y)\phi(s,y)\, \mathrm{d}s\, \mathrm{d}y\,, \tag{30}$$

where T' is the fundamental solution of (5) considered with the adjoint operator $\mathcal{P}'_\theta(\partial)$. One can verify that if $t > 0$ then $T'(t,x) = 0$. Thus combining (29) and (30) we conclude that $R_{\theta,\sigma}$ satisfies the equation

$$\mathcal{P}_\theta(\partial)R_{\theta,\sigma}(t,x) = 0 \tag{31}$$

for all $t > 0$ and $x \in \mathbb{R}^d$. The function $R_{\theta,\sigma}$ must be symmetric around zero. Hence $R_{\theta,\sigma}(t,x) = R_{\theta,\sigma}(-t,x)$ if $t < 0$. The value $R_{\theta,\sigma}(0,x) = \rho_{\theta,\sigma}(x)$ follows from the requirement of stationarity. See [12] (M o h a p l, 1994). The above argument can be reversed. That is, if $R_{\theta,\sigma}$ satisfies (23) and (24) then it satisfies (28) and serves as a proper representation of $\Gamma_{\theta,\sigma}$. Under our assumptions

$$R_{\theta,\sigma}(t,x) = -\frac{\sigma^2}{(2\pi)^d}\int_{\mathbb{R}^d} \frac{|\mathcal{B}_\theta(\mathrm{i}\,\lambda)|^2}{\mathcal{A}_\theta(\mathrm{i}\lambda)}\, \mathrm{e}^{\mathrm{i}\langle x,\lambda\rangle + t\mathcal{A}_\theta(\mathrm{i}\,\lambda)}\, \mathrm{d}\lambda \tag{32}$$

is a well defined solution of (23) and (24). The proposition is thus proved. □

There is an obvious resemblance between (23) and the Yule-Walker equations for an ordinary AR time series.

4. Estimation of θ and σ

If we know the covariance operator of the process and consider, for example, normally distributed observations, then it is not difficult to construct the likelihood function for θ and σ. Assume that c is a zero mean Gaussian random element and that we observe the values of

$$c(\phi_1), \dots, c(\phi_N)$$

for $\{\phi_1, \dots, \phi_N\} \subset \mathcal{S}$. Then (1) can be used to determine the covariance matrix

$$E_{\theta,\sigma} c(\phi_n)c(\phi_m) = (\Gamma_{\theta,\sigma}\phi_n)(\phi_m)\,, \tag{33}$$

$m, n = 1, \dots, N$, and the parameters θ and σ may be estimated via maximum likelihood or other known methods. The procedure outlined raises the question how to obtain an observation $c(\phi)$ for a given ϕ. When (5) has a solution in the sense of Definition 1 and continuous observations of ξ over an area $\Omega \subset \mathbb{R}^d$ are available then $c(\phi)$ may be determined from the definition: $c(\phi) = \int\limits_{\Omega} \phi(x)\xi(x)\,\mathrm{d}x$ for every test function with support in Ω. Recall that the support of ϕ consists of all $x \in \mathbb{R}^d$ such that $\phi(x) \neq 0$.

If $\hat{\theta}$ and $\hat{\sigma}$ are the maximum likelihood estimates and

$$\hat{\phi}_1 = \mathcal{P}_{\hat{\theta}}(\partial)\phi_1, \dots, \hat{\phi}_N = \mathcal{P}_{\hat{\theta}}(\partial)\phi_N$$

belong to $\mathcal{S}$ then we can use (6) to analyse the residuals

$$c(\hat{\phi}_n) = \hat{\sigma} W(\phi_n),$$

$n = 1, \dots, N$. The example below suggests that a suitable choice of test functions may substantially simplify the model and data analysis and leads to an interesting modification of the outlined method.

EXAMPLE 1. Consider the equation

$$(\partial_t - \theta\partial_x^2)\xi(t,x)\,\mathrm{d}t\,\mathrm{d}x = Z(\mathrm{d}t, \mathrm{d}x), \tag{34}$$

where $x \in \mathbb{R}$. To estimate $\theta \in (0,\infty)$ take an arbitrary $\phi \in \mathcal{S}$ which depends only on x so that the support of ϕ and of its derivatives is contained in $(0,1)$. If we multiply both sides of (34) by ϕ and integrate over x then (34) may be rewritten in the form

$$c_t(\phi) - c_0(\phi) = \theta \int\limits_0^t c_s(\partial_x^2\phi)\,\mathrm{d}s + W_t(\phi), \tag{35}$$

where $c_t(\phi) = \int\limits_0^1 \phi(x)\xi(t,x)\,\mathrm{d}x$ and $W_t(\phi) = \int\limits_0^t\int\limits_0^1 \phi(x)\,Z(\mathrm{d}s,\mathrm{d}x)$. Here Z is a Gaussian orthogonal measure with $E\big(Z(\mathrm{d}t,\mathrm{d}x)\big)^2 = \mathrm{d}t\,\mathrm{d}x$.

Suppose we fix $t > 0$ and observe $\xi(t,x)$ for all $x \in (0,1)$. Then θ can be determined almost surely. This is because $c_t(\phi)$ can be computed for every integrable function ϕ with support in $(0,1)$. In particular, for every function e_n defined by the relation $e_n(x) = \sqrt{2}\sin n\pi x$, $n = 1,2,\dots$, for $x \in (0,1)$ and $e_n(x) = 0$ otherwise. The random processes $w_n(t) = W_t(e_n)$, $n = 1,2,\dots$, are independent standard Wiener processes and the processes $c_n(e_n) = c_t(e_n)$ are ordinary independent Ornstein-Uhlenbeck processes determined by the equations

$$c_n(t) - c_n(0) = -\theta n^2\pi^2 \int\limits_0^t c_n(s)\,\mathrm{d}s + w_n(t). \tag{36}$$

The independent Gaussian observations $c_n(e_n)$ determine an asymptotically unbiased strongly consistent and efficient MLE estimate of θ.

The next theorem specifies the class of parabolic type equations that admit a determination of the parameter θ almost surely given space-continuous observations of the process ξ. The white noise W_t, $t > 0$, is a distribution valued Gaussian stochastic process generalizing ordinary Brownian motion in the sense that $EW_t(\phi)W_s(\psi) = (t \wedge s) \int_{\mathbb{R}^d} \phi(x)\psi(x)\, dx$ for every pair $t, s > 0$ and $\phi, \psi \in \mathcal{S}$.

THEOREM 4. ([4] Huebner, Rozovskii, 1995) *Suppose that equation* (1) *may be represented in the form*

$$c_t(\phi) = \int_0^t \big(\mathcal{P}_0(\partial_x) + \theta \mathcal{P}_1(\partial_x)\big) c_s(\phi)\, ds + W_t(\phi), \tag{37}$$

where $\mathcal{P}_0(\partial_x)$ and $\mathcal{P}_1(\partial_x)$ are commuting self-adjoint differential operators of order m_0 and m_1, respectively, and $\max(m_0, m_1) = 2m$ *for some natural number m. The notation ∂_x emphasizes that the operators do not act on the temporal variable t. Let the operators $\mathcal{P}_\theta(\partial_x) = \mathcal{P}_0(\partial_x) + \theta \mathcal{P}_1(\partial_x)$ be strongly elliptic with a complete orthonormal system of eigenvectors e_n. Then the following conditions are equivalent:*

i) *Order of $\mathcal{P}_1(\partial_x) \geq \big($order of $\mathcal{P}_\theta(\partial_x) - d\big)/2$, where d is the dimension of the x domain.*
ii) *The MLE's of θ are strongly consistent.*
iii) *The probability distributions of c for different θ are mutually singular.*

The previous theorem requires only space-continuous observations sampled at only one fixed time point. The spatial domain on which the process evolves is bounded. For the case of continuous temporal-spatial observations the parabolic type equations exhibit a more regular behavior.

THEOREM 5. ([12] Mohapl, 1994) *Let* (1) *have the form*

$$c_t(\phi) = \int_0^t \mathcal{P}_\theta(\partial_x) c_s(\phi)\, ds + W_t(\phi), \tag{38}$$

where the linear operator $\mathcal{P}_\theta(\partial_x)$ is defined by

$$\mathcal{P}_\theta(\partial_x) = \sum_{|k| \leq p} \theta_k \partial_k . \tag{39}$$

The last relation defines operator (4) *with $a_k(\theta) = \theta_k$ for all k. Suppose that W_t is as in the previous theorem and $\mathcal{P}_\theta(\partial_x)$ is invertible. Set*

$$R_\theta = \big(\mathcal{P}'_\theta(\partial_x) + \mathcal{P}_\theta(\partial_x)\big)^{-1} \tag{40}$$

and consider test functions $\phi_1, \ldots, \phi_N$ *which depend only on the spatial variable* $x \in \mathbb{R}^d$. *If* θ_0 *is the true value of* θ *and the matrix* M *with components* $m_{k,l} = \sum_n \int_{\mathbb{R}^d} \partial_k \phi_n(x) R_{\theta_0} \partial_l \phi_n(x)\, \mathrm{d}x$ *is invertible then for sufficiently large* $t > 0$, *the maximum likelihood estimators* $\theta(t)$ *of* θ_0 *are well defined,*

$$\lim_{t\to\infty} \theta(t) = \theta_0 \quad a.s.$$

and

$$\lim_{t\to\infty} \sqrt{t}\big(\theta(t) - \theta_0\big) = N(0, M^{-1})$$

in distribution.

Details of the construction of likelihood functions used in the above theorems are provided in the corresponding papers. For further generalizations see [15] (Piterbarg and Rozovskii, 1996).

The rest of this section deals with parameter estimation when only discrete observations are available. Suppose we have the heuristic model (8) leading to a generalized solution c. According to Theorem 3 we can identify our observations with values of a process ξ with covariance function determined by the covariance operator of c. Discretely observed values of the process ξ are not enough for calculation of any value $c(\phi)$, where ϕ is a prescribed test function. Thus we must rely only on the covariance function determined by c and specified up to some unknown parameters. The standard methods for estimation of these parameters are the maximum likelihood, quasi-likelihood, smoothed periodogram estimating equations etc. Their application is outlined in the following examples.

EXAMPLE 2. The parabolic stochastic differential equation

$$(\partial_t - \theta_1 \partial_x^2 + \theta_2)\xi(t,x)\, \mathrm{d}t\, \mathrm{d}x = \sigma Z(\mathrm{d}t, \mathrm{d}x) \tag{41}$$

with a one dimensional spatial variable is used to describe transport through a medium, the propagation of heat etc. If Z is the Gaussian orthogonal white noise measure then its space-time stationary solution can be represented for $t \geq s > 0$ by the process

$$\xi(t,x) = \int_{-\infty}^{\infty} T(t-s, x-y)\xi(s,y)\, \mathrm{d}y + \sigma \int_s^t \int_{-\infty}^{\infty} T(t-u, x-y)\, Z(\mathrm{d}u, \mathrm{d}y)\,, \tag{42}$$

where the function T is defined by the relation

$$T(t,x) = \frac{1}{\sqrt{4\pi\theta_1 t}}\, \mathrm{e}^{-x^2/(4\theta_1 t)}\, \mathrm{e}^{-\theta_2 t} \tag{43}$$

if $t > 0$ and is zero otherwise. The process $\big(\xi(t,x)\big)_{x\in(-\infty,\infty)}$ is an Ornstein-Uhlenbeck for every $s \geq 0$. The covariance function of ξ is, according to Proposition 2,

$$R(t,x) = \frac{1}{\sqrt{4\pi\theta_1 t}} \int_{-\infty}^{\infty} \exp\left\{-\frac{(x-y)^2}{4\theta_1 t}\right\} \exp\{-\theta_2 t\}\rho(y)\,\mathrm{d}y \tag{44}$$

for all $t \geq 0$, where

$$\rho(y) = \frac{\sigma^2}{2\sqrt{\theta_2\theta_1}} \exp\left\{-|y|\sqrt{\theta_2/\theta_1}\right\}, \tag{45}$$

t is the time and x the space variable. For fixed x, equation (42) is an ordinary Itô process. Hence the parameters may be estimated by means of methods developed for ordinary Itô equations. See for example [3] (Florens-Zmirou, 1993).

Another possibility is to use the representation (42) directly. Suppose we have observations $\{\xi(t_k,x_n),\ k = 1,\dots,K,\ n = 1,\dots,N\}$. Denote by $\xi(t_k,\cdot)$ the column vector $\big(\xi(t_k,x_1),\dots,\xi(t_k,x_N)\big)^T$ and by $\Gamma_\tau(\theta,\sigma)$ the matrix $\Big(E_{\theta,\sigma}\big(\xi(\tau,x_n)\xi(0,x_m)\big)\Big)_{n,m=1}^{N}$. The first integral in (42) denotes the conditional expectation of the value $\xi(t,x)$ conditioned on the continuously observed path $\{\xi(s,y) : y \in (-\infty,\infty)\}$. Since only discrete observations are available, we must replace it by the conditional expectation based on the pointwise observations:

$$\Big(E_{\theta,\sigma}\big[\xi(t_k,x_n) \mid \xi(t_{k-1},\cdot)\big]\Big)_{n=1}^{N} = \Gamma_{t_k-t_{k-1}}(\theta,\sigma)\Gamma_0^{-1}(\theta,\sigma)\xi(t_{k-1},\cdot). \tag{46}$$

The Itô integral in (42) provides the conditional variance of what we observe at time t conditioned on what we have seen in time $s < t$. The conditional variance may be evaluated:

$$\mathrm{var}_{\theta,\sigma}\big(\xi(t_k,\cdot) \mid \xi(t_{k-1},\cdot)\big) = \Gamma_0(\theta,\sigma) - \Gamma_{2(t_k-t_{k-1})}(\theta,\sigma). \tag{47}$$

Hence, for the true values of θ and σ, the vectors

$$\big(\Gamma_0(\theta,\sigma) - \Gamma_{2(t_k-t_{k-1})}(\theta,\sigma)\big)^{-1/2}\big(\xi(t_k,\cdot) - \Gamma_{t_k-t_{k-1}}(\theta,\sigma)\Gamma_0^{-1}(\theta,\sigma)\xi(t_{k-1},\cdot)\big), \tag{48}$$

$k = 2,\dots,N$, are approximately multivariate normal with zero mean and unit covariance. This can be successfully used for construction of a quasi-likelihood function and subsequent estimation of θ and σ. For details see [9], [10] (Mohapl, 1998).

EXAMPLE 3. The elliptic stochastic differential equation with a two dimensional spatial variable is given by the relation

$$(\partial^2_{x_1} + \partial^2_{x_2} - \theta^2)\xi(x)\,\mathrm{d}x = \sigma Z(\mathrm{d}x)\,. \tag{49}$$

It describes, among other things, the displacement of a membrane of a homogeneous isotropic material, stationary heat conduction etc. The stationary solution can be represented in the form

$$\xi(x) = -\frac{\sigma}{2\pi}\int\limits_{\mathbb{R}^2} K_0(-|x-y|\theta)\; Z(\mathrm{d}y)\,, \tag{50}$$

where K_0 is the modified Bessel function of second type and order zero, and $|x|$ is the Euclidean norm of x. The covariance function of the process (50) is

$$R(x) = \frac{\sigma^2}{4\pi\theta}|x|K_1\big(|x|\theta\big)\,, \tag{51}$$

where K_1 is the modified Bessel function of second type and order one. A direct likelihood analysis of random fields like this one is in detail e.g. in [6] (Jones and Vecchia, 1993).

Another possibility is suggested in [9], [10] (Mohapl, 1998). We assume availability of continuous observations, derive a conditional likelihood function for the continuously observed trajectory and then discretize. The resulting quasi-likelihood estimating equations for θ and $\sigma > 0$ based on a set $\{\xi(x_1),\dots,\xi(x_N)\}$ of discrete observations are

$$\frac{1}{N}\sum_{n=1}^{N}\xi^2(x_n) - \sigma^2 R_\theta(0) = 0\,, \tag{52}$$

$$\sum_{n=1}^{N}\sum_{m=1}^{N} k(x_n,x_m)\big(\xi(x_n)\xi(x_m) - \sigma^2 R_\theta(x_n - x_m)\big) = 0\,, \tag{53}$$

where k is a smoothing kernel. An optimal k may, for example, minimize the asymptotic variance of the estimators θ and σ. Reasonable results are obtained for example by choosing $k(x,y) = R_1(x-y)$. To obtain an idea about variability of the resulting estimator we can use for example independent simulations.

EXAMPLE 4. The hyperbolic equation with one dimensional time and space variables is given by

$$\partial^2_{tx}c + \theta_2\partial_t c + \theta_1\partial_x c + \theta_1\theta_2 c = \sigma W\,. \tag{54}$$

If $W = 0$ then its canonical form describes, for example, propagation of a wave along a water surface. For $t > 0$ and $x > 0$ we can represent the stationary solution of equation (54) as

$$\xi(t,x) = q_1\,\mathrm{e}^{-\theta_1 t} + q_2\,\mathrm{e}^{-\theta_2 x} + q_3\,\mathrm{e}^{-\theta_1 t - \theta_2 x} + \sigma \int_0^x \int_0^y \mathrm{e}^{-\theta_1(t-s)-\theta_2(x-y)}\ Z(\mathrm{d}s, \mathrm{d}y)\,, \tag{55}$$

where the q's are independent, identically distributed (i.i.d.) random variables with zero mean and variance $\sigma^2/4\theta_1\theta_2$. The representation (55) of the solution to (54) has covariance function

$$R_{\theta,\sigma}(t,x) = \frac{\sigma^2}{4\theta_1\theta_2}\,\mathrm{e}^{-\theta_1|t|-\theta_2|x|}\,. \tag{56}$$

Two cases may be considered when dealing with discrete observations from time-space continuous processes. We may either keep equal spacing between the observations and increase their number by broadening the monitored area, or we may keep the area fixed but increase density of our observations. The following theorem investigates these two possibilities for the process (55).

Let $\big(\xi(t_n, x_m)\big)_{n=0,m=0}^{N,M}$ be observations of the process (55) with true parameters θ_{01} and θ_{02}. We take $t_n = t_0 + nh$, $x_m = x_0 + mk$ so that the points (t_n, x_m) form a rectangular net with nodes at a fixed distance $h = a/N$ and $k = b/M$. We can estimate the parameters either using the likelihood obtained by *conditioning* on the boundary elements or by the *unconditional* likelihood function.

THEOREM 6. *If $\hat{\theta}_1$, $\hat{\theta}_2$ are obtained either by the maximization of the conditional or unconditional likelihood function then the estimators are strongly consistent.*

i) *For an unbounded area, the vector $\sqrt{NM}(\theta_{01} - \hat{\theta}_1, \theta_{02} - \hat{\theta}_2)^T$ is asymptotically normal with zero mean and covariance matrix*

$$\Gamma = \mathrm{diag}\big((\mathrm{e}^{2\theta_{01}h} - 1)/h^2, (\mathrm{e}^{2\theta_{02}k} - 1)/k^2\big)\,.$$

ii) *For a bounded area, the conditional likelihood provides estimators such that the vector $\big(\sqrt{M}(\hat{\theta}_1 - \theta_{01}), \sqrt{N}(\hat{\theta}_2 - \theta_{02})\big)^T$ is, for increasing N and M, asymptotically normal with covariance matrix*

$$\Gamma = 2\,\mathrm{diag}(\theta_{01}/a, \theta_{02}/b)\,.$$

iii) *For a bounded area, the unconditional likelihood yields estimators such that the vector $\big(\sqrt{M}(\hat{\theta}_1 - \theta_{01}), \sqrt{N}(\hat{\theta}_2 - \theta_{02})\big)^T$ has, for increasing N and M, an asymptotic normal distribution with covariance matrix*

$$\Gamma = 2\,\mathrm{diag}\big(\theta_{01}^2/(1 + a\theta_{01}), \theta_{02}^2/(1 + b\theta_{02})\big)\,.$$

P r o o f.

i) The observations form a doubly geometric series. The result is thus a consequence of the proposition in [8; Section 3] (M a r t i n, 1990).

For the proof of ii) and iii) see [11] (M o h a p l, 1997). □

5. Representation of c

The results outlined in Section 4 required the existence of an observable representation of the generalized random field c solving equation (5). We have already mentioned that the class of SPDE's with the representation $c(\phi) = \int_{\mathbb{R}^d} \phi(x)\xi(x)\,\mathrm{d}x$, i.e., with a solution in the sense of Definition 1, is fairly narrow.

THEOREM 7. *For every distribution-valued random element c there exists a random field η with one and only one spectral measure F such that*

$$c(\phi) = \int_{\mathbb{R}^d} \phi(x)\eta(x)\,F(\mathrm{d}x) \tag{57}$$

for every rapidly decreasing $\phi \in \mathcal{S}$.

P r o o f. See [18] (Y a g l o m, 1957). □

A spectral measure F is slowly growing if $\int_{\mathbb{R}^d} (1+|x|^2)^{-m}\,F(\mathrm{d}x) < \infty$ for some natural number m and the measure is σ-finite. We use Theorem 7 as follows. We start with the model (5), derive the abstract equation (1) and compute the generalized random field c. Then we apply Theorem 7 and obtain the process η satisfying (57). If $F(\mathrm{d}x) = \mathrm{d}x$ then $\xi = \eta$ is the desired solution to (5). Otherwise we have to modify the model (5) and replace it by

$$\mathcal{P}_\theta(\partial)\eta(x)F(\mathrm{d}x) = \sigma Z(\mathrm{d}x)\,. \tag{58}$$

In the case when $F(\mathrm{d}x) = \rho(x)\,\mathrm{d}x$, where ρ is a strictly positive "slowly growing" function on $\mathbb{R}^d$, it is possible to divide both sides of (58) by ρ. Measure $Z(\mathrm{d}x)$ in equation (5) is thus replaced by $\rho^{-1}(x)Z(\mathrm{d}x)$. This means that we impose a new assumption on the covariance structure of the noise in the model (5).

The abstract theory of SPDE's introduces the above mentioned change in covariance structure in a less obvious manner through the choice of a system of inner products on the space of test functions and their duals. Since the importance of the choice of the inner-product for the interpretation of the original model (5) is not very often emphasized in the literature, we demonstrate it by the following example.

EXAMPLE 5. Let us consider the heat equation

$$\Big(\partial_t - \sum_{k=1}^{d} \partial^2_{x_k} + \theta^2\Big)\xi(t,x)\,\mathrm{d}t\,\mathrm{d}x = Z(\mathrm{d}t,\mathrm{d}x)\,, \tag{59}$$

where Z is a Gaussian orthogonal measure, $E\big(Z(\mathrm{d}t,\mathrm{d}x)\big)^2 = \mathrm{d}t\,\mathrm{d}x$, and $\theta > 0$. We formalize the equation which leads to the problem

$$c_t(\phi) - c_0(\phi) = \int_0^t c_s\big(\mathcal{P}_\theta(\partial_x)\phi\big)\ \mathrm{d}s + W_t(\phi) \tag{60}$$

where $\mathcal{P}_\theta(\partial_x) = \sum_{k=1}^{d} \partial^2_{x_k} - \theta^2$ and $W_t(\phi) = \int_0^t \int_{\mathbb{R}^d} \phi(x)\ Z(\mathrm{d}s,\mathrm{d}x)$ if $t > 0$. We consider ϕ in the space $\mathcal{S}$ of rapidly decreasing functions on $\mathbb{R}^d$ with the Schwartz topology. The space $\mathcal{S}$ becomes a nuclear space using the system of inner products

$$(\phi,\psi)_m = \int_{\mathbb{R}^d} \hat\phi(\zeta)\hat\psi^*(\zeta)(|\zeta|^2+1)^m\ \mathrm{d}\zeta\,, \tag{61}$$

$m = 0, 1, \dots$. The hat denotes the Fourier transform of the original function. Using the mth product we can complete $\mathcal{S}$ into a Hilbert space $H_m, (\cdot\,,\cdot)_m$. For $m = 0$ the index is usually omitted. Details of the construction, the properties and the use of such spaces are given, for example, in [5] (I t ô, 1984).

If m is sufficiently large then W_t may be thought of as an H_{-m} valued random element possessing an H_m-valued representation ω_t. We will seek c_t in the form $c_t(\phi) = (\xi_t, \phi)_m$, where ξ_t is an H_m-valued random process. In other words, we wish to solve the equation

$$(\xi_t,\phi)_m - (\xi_0,\phi)_m = \int_0^t \big(\xi_s, \mathcal{P}_\theta(\partial_x)\phi\big)_m\ \mathrm{d}s + (\omega_t,\phi)_m\,. \tag{62}$$

Notice that m must not only be large enough to admit W to have an H_m valued modification but that we also need ϕ in the domain of $\mathcal{P}_\theta(\partial_x)$ and $\mathcal{P}_\theta(\partial_x)\phi \in H_m$ for every $\phi \in H_m$. Its solution ξ_t is supposed to satisfy (60) (hence (62)) almost surely for each $\phi \in H_m$. The process ω_t admits a modification with almost surely continuous trajectories with respect to t (see e.g. [16] (W a l s h, 1986)) and has a nuclear covariance operator, because

$$E(\omega_t,\phi)(\omega_t,\psi) = \min(t,s)\int_{\mathbb{R}^d}\Big(\sum_{k=1}^{d}\partial^2_{x_k} - 1\Big)^{-m}\phi(x)\psi(x)\ \mathrm{d}x\,.$$

If the process ξ_t is well defined then the equation (62) can be rewritten as

$$\begin{aligned}&\int_{\mathbb{R}^d} \hat{\xi}_t(\zeta)\hat{\phi}(\zeta)\rho_m(\zeta)\,\mathrm{d}\zeta - \int_{\mathbb{R}^d} \hat{\xi}_0(\zeta)\hat{\phi}(\zeta)\rho_m(\zeta)\,\mathrm{d}\zeta \\ &= \int_0^t \int_{\mathbb{R}^d} \hat{\xi}_s(\zeta)\hat{\phi}_m\big(|\zeta|^2+\theta^2\big)\rho_m(\zeta)\,\mathrm{d}\zeta\,\mathrm{d}s + \int_{\mathbb{R}^d} \hat{\omega}_t(\zeta)\hat{\phi}(\zeta)\rho_m(\zeta)\,\mathrm{d}\zeta\end{aligned} \tag{63}$$

arising by substitution from (61). The weight function ρ_m is given by $\rho_m(\zeta) = \big(|\zeta|^2+1\big)^m$. However, this means that $\hat{\xi}_t$ satisfies the equation

$$\hat{\xi}_t(\zeta) - \hat{\xi}_t(\zeta) = \big(|\zeta|^2+\theta^2\big)\int_0^t \hat{\xi}_s(\zeta)\,\mathrm{d}s + \hat{\omega}_t(\zeta) \tag{64}$$

almost surely for almost each ζ. The last equation has solution

$$\hat{\xi}_t(\zeta) = \mathrm{e}^{-(|\zeta|^2+\theta^2)t}\,\hat{\xi}_0(\zeta) + \int_0^t \mathrm{e}^{-(|\zeta|^2+\theta^2)(t-s)}\,\mathrm{d}\hat{\omega}_t(\zeta)\,. \tag{65}$$

The desired solution $\xi_t(x)$ can be obtained by the inverse Fourier transformation of $\hat{\xi}_t(\zeta)$. However, it remains to be proved that the expression obtained in this way is an H_m valued process. The existence of ξ_t is proved for a much more general class of parabolic equations in [2] (Curtain and Falb, 1971).

If ξ_t is stationary in t then the initial value $\xi_0(x)$ can be obtained as follows. Combining the left hand side of (63) with (65) one may derive for large t and s, $t > s$, the approximate equality

$$\begin{aligned}&E(\xi_t,\phi)(\xi_s,\psi) \\ &= \int_{\mathbb{R}^d} \frac{1}{2(|\zeta|^2+\theta^2)}\Big(\mathrm{e}^{-(|\zeta|^2+\theta^2)(t-s)} - \mathrm{e}^{-(|\zeta|^2+\theta^2)(t+s)}\Big)\hat{\phi}(\zeta)\hat{\psi}(\zeta)\rho_m^{-1}(\zeta)\,\mathrm{d}\zeta\,.\end{aligned} \tag{66}$$

For $t = s$ and t approaching infinity

$$E(\xi_\infty,\phi)(\xi_\infty,\psi) = \frac{1}{2}\int_{\mathbb{R}^d} \frac{1}{(|\zeta|^2+\theta^2)}\hat{\phi}(\zeta)\hat{\psi}(\zeta)\rho_m^{-1}(\zeta)\,\mathrm{d}\zeta\,. \tag{67}$$

Consequently, the random field $\xi_0(x)$ has the same Gauss distribution as the element $\eta(x) = \int_{\mathbb{R}^d} \mathrm{e}^{-\mathrm{i}\langle x,\zeta\rangle}\,M(\mathrm{d}\zeta)$ defined by the Gauss orthogonal measure M with spectral density

$$f(\zeta) = \frac{1}{2}\int_{\mathbb{R}^d} \frac{1}{(|\zeta|^2+\theta^2)(|\zeta|^2+1)^m}\,\mathrm{d}\zeta\,. \tag{68}$$

As outlined in [12] (Mohapl, 1994),

$$E c_{\infty}(\phi) c_{\infty}(\psi) = \frac{1}{2} \int_{\mathbb{R}^d} \frac{1}{(|\zeta|^2 + \theta^2)} \hat{\phi}(\zeta) \hat{\psi}(\zeta) \, \mathrm{d}\zeta \,. \tag{69}$$

This means that the covariance of ξ_t is damped by the weight ρ_m^{-1} implementing the nuclearity into our calculations.

At first glance, application of Hilbert spaces seems to be a convenient way to well behaved representations. The trouble is that processes with differentiable trajectories and a prescribed rate of decay when the argument approaches infinity are not of much interest. If such a process is simulated, the trajectories often appear as deterministic functions. If the experimenter has only one such trajectory, and this is often the case, he or she will reach rather for deterministic methods of data analysis.

Suppose that we derive the heuristic model (59) with $Z(\mathrm{d}t, \mathrm{d}x)$ replaced by $\mathrm{d}\omega_t(x)\,\mathrm{d}x$ (the differential applies to the variable t). Then (62) may be understood as the mathematical definition of (59). This fits well the procedure in Section 3. We just use the inner product $(\cdot\,,\cdot)_m$ instead of the ordinary Lebesgue integral. However, the assumption that the noise has such smooth trajectories (in H_m) does not seem realistic. The abstract solution $c_t(\phi) = (\xi_t, \phi)_m$, whose existence was proved by finding ξ_t, can still be interpreted in terms of Theorem 7 above.

6. Summary

The Schwartz distribution approach to the SPDE provides:

1) A mathematically exact definition of the SPDE and its solution.
2) A precise relation between the equation, its solution and observations which therefore contributes to the correct interpretation of data.
3) Criteria for the choice of an orthogonal measure Z with covariance structure that makes solution of the equation mathematically feasible.
4) Methods for computation of the solution.
5) Tools for determining the covariance structure of the solution.
6) The use of the SPDE for estimation and goodness-of-fit assessment.
7) A certain degree of caution is needed when interpreting the resulting process.

REFERENCES

[1] AL-GWAIZ, M. A.: *Theory of Distributions*, Marcel Dekker, Inc., New York-Basel-Hong Kong, 1992.

[2] CURTAIN, R. F.—FALB, P. L.: *Stochastic differential equations in Hilbert space*, J. Differential Equations **10** (1971), 412–130.

[3] FLORENS-ZMIROU, D.: *On estimating the diffusion coefficient from discrete observations*, J. Appl. Probab. **30** (1993), 790–804.

[4] HUEBNER, M.—ROZOVSKII, B. L.: *On asymptotic properties of maximum likelihood estimators for parabolic stochastic PDE's*, Probab. Theory Related Fields **103** (1995), 143–163.

[5] ITÔ, K.: *Foundations of Stochastic Differential Equations in Infinite Dimensional Spaces*, SIAM, Philadelphia, Pennsylvania, 1984.

[6] JONES, R. H.—VECCHIA, A. V.: *Fitting Continuous ARMA models to unequally spaced spatial data*, J. Amer. Statist. Assoc. **88** (1993), 947–954.

[7] LIPTSER, R. S.—SHIRYAYEV, A. N.: *Statistics of Random Processes I. General Theory* (2nd ed., 1st ed. 1977), Springer-Verlag, New York-Berlin-Heidelberg-Tokyo, 1984.

[8] MARTIN, R. J.: *The use of time-series models and methods in the analysis of agricultural field trials*, Comm. Statist. Theory Methods **19** (1990), 55–81.

[9] MOHAPL, J.: *A stochastic advection-diffusion model for the rocky flats soil plutonium data*, Ann. Inst. Statist. Math. (1999) (To appear).

[10] MOHAPL, J.: *Discrete sample estimation for Gaussian random fields generated by stochastic partial differential equations*, Comm. Statist. Stochastic Models **14** (1998), 883–903.

[11] MOHAPL, J.: *On estimation in the planar Ornstein-Uhlenbeck process*, Comm. Statist. Stochastic Models **13** (1997), 435–455.

[12] MOHAPL, J.: *Maximum likelihood estimation in linear infinite dimensional models*, Comm. Statist. Stochastic Models **10** (1994), 781–794.

[13] NAMACHCHIVAYA, N. S.: *Stochastic Structural Dynamics. Proceedings of the Symposium held at the University of Illinois at Urbana Champaign October 30. – November 1, 1988* (N. S. Namachchivaya, H. H. Hilton, Y. K. Wen, eds.), University of Illinois at Urbana Champaign, Urbana, Illinois, 1989.

[14] PARTHASARATHY, K. R.: *Introduction to Probability and Measure*, MacMillan Co., India, 1977.

[15] PIETERBARG, L.—ROZOVSKII, B.: *Estimating unknown parameters in SPDE's under discrete observations in time.* Preprint, 1996.

[16] WALSH, J. B.: *An introduction to differential equations.* In: Lecture Notes in Math. 1180, Springer Verlag, Berlin-New York-Heidelberg, 1986, pp. 266–437.

[17] WHITTLE, P.: *Topographic correlation, power law covariance functions and diffusion*, Biometrika **49** (1962), 305–314.

[18] YAGLOM, A. M.: *Some classes of random fields in n-dimensional space related to stationary random processes*, Theory Probab. Appl. **3** (1957), 273–320.

[19] YOSIDA, K.: *Functional Analysis* (4th ed.), Springer-Verlag, New York-Berlin-Heidelberg, 1974.

Received August 14, 1995
Revised September 2, 1996

Department of Statistics
and Actuarial Science
University of Waterloo
Waterloo N2L 3G1
Ontario
CANADA
E-mail: jmmohapl@jeeves.uwaterloo.ca