Proceedings of Machine Learning Research 103:314–323, 2019 ISIPTA 2019

Random Set Solutions to Stochastic Wave Equations

Michael Oberguggenberger MICHAEL.OBERGUGGENBERGER@UIBK.AC.AT
Unit of Engineering Mathematics, Innsbruck University, Austria

Lukas Wurzer WURZER.LUKAS@GMAIL.COM
Johannesplatz 1, 77815 Bühl, Germany

Abstract

This paper is devoted to three topics. First, to prove a measurability theorem for multifunctions with values in non-metrizable spaces, which is required to show that solutions to stochastic wave equations with interval parameters are random sets; second, to apply the theorem to wave equations in any space dimension; and third, to compute upper and lower probabilities of the values of the solution in the case of one space dimension.

Keywords: random sets, fundamental measurability theorem, non-metrizable spaces, stochastic wave equations

1. Introduction

In order to motivate the subject of the paper, consider a stochastic dynamical system

$$\frac{\mathrm{d}}{\mathrm{d}t}u(t) = F\big(u(t),\lambda\big) + G\big(u(t),\lambda\big)N(t) \tag{1}$$

where $u(t)$ is the response of the system, λ comprises the system parameters, and $N(t)$ is an external excitation in the form of a stochastic process. The response $u(t)$ may be scalar-valued or n-dimensional, the parameters λ take values in $\mathbb{R}^m$, say. For example, $u(t)$ may be the response of a building under earthquake excitation, the vibrational response of a bridge under random traffic, a biological quantity evolving under external disturbances, or – in financial mathematics – the value of an option under the uncertain development of the price of the underlying asset. The standard model for the noise process $N(t)$ is Gaussian white noise as the (generalized) derivative of Brownian motion. In this case, (1) can be interpreted as an Itô differential equation.

We need to introduce a more specific notation. The noise is a stochastic process on some time interal T and a probability space (Ω,Σ,P). Accordingly, the response is also a stochastic process, i.e., a map $T\times\Omega\to\mathbb{R}^n$, which is measurable with respect to ω at each fixed t. In addition, the response depends on the values λ of the system parameters. Thus we will write more accurately $u_\lambda(t,\omega)$ in place of $u(t)$ from now on.

Next, assume that the system parameters λ are uncertain and described by some imprecise probability model. We put ourselves in the framework of random sets [4, 14].

To recall the notion of a random set, let S be a topological space, $\mathscr{P}(S)$ the power set of S and $\mathscr{B}(S)$ the Borel σ-algebra on S. A *random set* on a probability space (Ω,Σ,P) with values in S is a multifunction $X:\Omega\to\mathscr{P}(S)$ such that the upper inverses

$$X^{-}(B) = \{\omega\in\Omega : X(\omega)\cap B\neq\emptyset\} \tag{2}$$

are measurable for every Borel set $B\in\mathscr{B}(S)$.

Let the system parameters in (1) be described by a random set $\Lambda:\Omega\to\mathbb{R}^m$. (Without restriction of generality, (Ω,Σ,P) may be taken to be the same probability space as the one underlying the stochastic system.) Then the response is also a multifunction, which can be viewed from various perspectives. First, at fixed $t\in T$, one may consider the multifunction

$$X(t,\omega) = \{u_\lambda(t,\omega) : \lambda\in\Lambda(\omega)\} \tag{3}$$

on Ω with values in $\mathbb{R}^n$. Second, and more generally, the response can be seen as a multifunction on Ω with values in the space of trajectories of system (1), usually the space $\mathscr{C}(T,\mathbb{R}^n)$ of continuous functions on T with values in $\mathbb{R}^n$, equipped with the topology of uniform convergence on compact subsets of T, i.e.,

$$X(\omega) = \{u_\lambda(\cdot,\omega) : \lambda\in\Lambda(\omega)\}. \tag{4}$$

In the simplest case when the random set reduces to an interval $\Lambda = [\underline{\lambda},\overline{\lambda}]$, this becomes

$$X(\omega) = \{u_\lambda(\cdot,\omega) : \underline{\lambda}\leq\lambda\leq\overline{\lambda}\}. \tag{5}$$

In order to go on with computing probability bounds, one should show that the multifunctions (3), (4), (5) are actually random sets.

The theory of random set valued solutions to (1) has been developed in [18, 19] and applied to problems of earthquake engineering [20] as well as geotechnical reliability [15, 16]. In both cases, the excitation is reasonably modelled by Gaussian white noise, whereas the uncertainty of system parameters such as damping coefficients or soil-related properties are more plausibly modelled by random sets. We mention that other set-valued theories for treating this issue have been proposed, such as set-valued stochastic integrals [11], mutational analysis [12], fuzzy stochastic processes [21]; for the relation of these approaches to stochastic differential inclusions, see [13].

An important ingredient in proving that the multifunctions (3), (4), (5) are random sets is the Fundamental Measurability Theorem [14], which can be applied here in its classical form for Polish spaces (see Section 2).

The situation is quite different when the independent variable in (1) is multi-dimensional and has a time component t as well as a spatial component x. Then (1) becomes a stochastic partial differential equation, and the trajectories of the solutions turn out to be generalized stochastic processes already in simplest cases. Random set solutions to stochastic partial differential equations have not been addressed in the literature. This paper undertakes first steps in this direction. *The* basic model of a stochastic partial differential equation [6] is the stochastic wave equation

$$\left(\partial_t^2 - c^2\Delta\right) u_c = \dot{W} \tag{6}$$

on $\mathbb{R}^{d+1}$, where $\dot{W}$ is space-time Gaussian white noise, $\Delta = \partial_{x_1}^2 + \cdots + \partial_{x_d}^2$ is the Laplacian, and $c > 0$ is a constant, the *propagation speed*. The underlying probability space is the *white noise probability space*, see Subsection 3.2. It is well-known [6, 25] that in space dimensions $d \geq 2$, the solution u_c is a generalized stochastic process, that is, its trajectories $u_c(\omega)$ belong to the Schwartz space of distributions $\mathscr{D}'(\mathbb{R}^{d+1})$. (For the convenience of the reader, all required notions from the theory of distributions are collected in Appendix B.)

Replacing the propagation speed c by an interval $[\underline{c}, \overline{c}]$ with $\underline{c} > 0$ results in the multifunction

$$X(\omega) = \{u_c(\omega) : \underline{c} \leq c \leq \overline{c}\} \tag{7}$$

on Ω with values in the power set of $\mathscr{D}'(\mathbb{R}^{d+1})$. Endowing $\mathscr{D}'(\mathbb{R}^{d+1})$ with the Borel σ-algebra generated by its weak topology (see Appendix A), one would like to show that X is a random set. However, $\mathscr{D}'(\mathbb{R}^{d+1})$ is not metrizable (see Appendix B) and so the Fundamental Measurability Theorem for Polish spaces does not apply.

The first part of the paper will be devoted to proving a generalization of the Fundamental Measurability Theorem, showing that a sequentially compact, Effros measurable multifunction (see Section 2) with values in the dual of a separable topological vector space, which is itself separable and a Souslin space, is a random set. The second part of the paper provides the details required to apply this result to the wave equation in arbitrary space dimensions, culminating in the proof that the multifunction X in (7) is a random set. In the third part we will focus on the stochastic wave equation in one space dimension ($d = 1$). In this case, the solution is a modified Brownian sheet and consequently it has almost surely continuous trajectories. Thus one can consider the $\mathbb{R}$-valued multifunctions

$$X(x,t,\omega) = \{u_c(x,t,\omega) : \underline{c} \leq c \leq \overline{c}\} \tag{8}$$

at fixed (x,t). It turns out that the maps $(r,\omega) \to \frac{2}{t} u_{1/r}(x,t,\omega)$ are Brownian motions with respect to $r \in [0,\infty)$. Using well-known formulas for the first hitting times, this will allow us to compute the upper and lower distribution functions of the random sets (8).

The paper ends with some conclusions and outlook for further research. The appendices collect the required notions from topology and the theory of distributions.

The presented results have been obtained by the second author in [26]. Various other generalizations of the Fundamental Measurability Theorem beyond Polish spaces can be found in [1, 4, 8].

2. Measurability of Multifunctions

As suggested by the defining property, a random set is also referred to as a *Borel measurable multifunction*. Various measurability properties will play a role in the Fundamental Measurability Theorem. A multifunction X is called *Effros measurable* if its upper inverses in (2) are measurable for every open set $B \subset S$. The multifunction X is called *graph measurable* if its graph

$$\operatorname{Graph}(X) = \{(\omega, x) \in \Omega \times S : x \in X(\omega)\}$$

is measurable, that is, belongs to the product σ-algebra $\Sigma \otimes \mathscr{B}(S)$. A random variable ξ with values in S is called a *selection* of X if $\xi(\omega) \in X(\omega)$ for almost all $\omega \in \Omega$. Let X be a closed-valued random set. A countable family of selections ξ_n is said to be a *Castaing representation* of X if

$$X(\omega) = \operatorname{cl}\{\xi_n(\omega), n \geq 1\} \tag{9}$$

for (almost) all $\omega \in \Omega$, where cl denotes the closure in S.

Remark 1 *(a) If a multifunction is Effros measurable, then its lower inverses*

$$X_-(B) = \{\omega \in \Omega : X(\omega) \subset B\} \tag{10}$$

are measurable for every closed set $B \subset X$, and vice versa. This follows from the formula $X^-(B^c) = (X_-(B))^c$.

(b) If a closed valued multifunction $X : \Omega \to S$ into some topological space S has a Castaing representation, then it is Effros measurable. Indeed, let $(\xi_n)_{n\in\mathbb{N}}$ be a Castaing representation of X and let $B \subset S$ open. On the one hand, $\xi_n(\omega) \in X(\omega)$ for (almost) all $\omega \in \Omega$ and all $n \in \mathbb{N}$. Thus $\{\omega : \xi_n(\omega) \in B\} \subset X^-(B)$ for all $n \in \mathbb{N}$. On the other hand, let $\omega \in X^-(B)$. Since B is open, (9) implies the existence of an $n \in \mathbb{N}$ such that $\xi_n(\omega) \in B$ for (almost) all $\omega \in \Omega$. Thus

$$X^-(B) = \bigcup_{n\in\mathbb{N}} \{\omega : \xi_n(\omega) \in B\}.$$

Measurability of $X^-(B)$ follows from measurability of the selections ξ_n.

(c) When talking about multifunctions $X : \Omega \to \mathscr{P}(S)$ we assume from now on that

$$X(\omega) \neq \emptyset \quad \text{for (almost) all } \omega \in \Omega.$$

Recall that a subset of a topological space S is called *dense* if its closure equals S. The topological space S is *separable* if it contains a dense countable subset. A subset of A of S is *sequentially compact* if every sequence in A contains a converging subsequence (with limit in A). Finally, a *Polish space* is a separable, metrizable and complete topological space, while a *Souslin space* is a Hausdorff space which is the continuous image of a Polish space.

Let (Ω, Σ, P) be a complete probability space, S a Polish space, and $X : \Omega \to \mathscr{P}(S)$ a closed-valued multifunction. The Fundamental Measurability Theorem [14] says that X is Borel measurable if and only if it is Effros measurable if and only if it is graph measurable if and only if it admits a Castaing representation.

In what follows, E is a Hausdorff topological vector space and E' its continuous dual, equipped with its weak topology (see Appendix A). We denote by $\langle x, e\rangle$ the action of $x \in E'$ on $e \in E$. The probability space (Ω, Σ, P) is assumed to be complete throughout. Here is the main result of this section.

Theorem 2 *Assume that both E and E' are separable and that E' is Souslin. Let $X : \Omega \to \mathscr{P}(E')$ be an Effros measurable multifunction with sequentially compact values. Then X is Borel measurable, i. e., a random set.*

The proof requires a few preparations. Let $(e_i)_{i\in\mathbb{N}} \subset E$ and $(e_i')_{i\in\mathbb{N}} \subset E'$ be dense countable subsets. Given $x \in E'$, the sets

$$U_{mn}(x) = \left\{y \in E' : |\langle y - x, e_1\rangle| < \tfrac{1}{m}, \dots, |\langle y - x, e_n\rangle| < \tfrac{1}{m}\right\}$$

with $m, n \in \mathbb{N}$, form a countable set of neighborhoods of x. In the following lemmas, E' is not required to be a Souslin space.

Lemma 3 *Let $A \subset E'$ be sequentially compact and $x_0 \in E'$ such that $x_0 \notin A$. Then there exists $\tilde{n} \in \mathbb{N}$ such that $U_{1\tilde{n}}(x_0) \cap A = \emptyset$.*

Proof We assume the converse and derive a contradiction. So suppose that $x_0 \notin A$ and $U_{1n}(x_0) \cap A \neq \emptyset$ for all $n \in \mathbb{N}$. Then one can find a sequence $(x_n)_{n\in\mathbb{N}}$ such that $x_n \in U_{1n}(x_0) \cap A$ for all $n \in \mathbb{N}$. Since $(x_n)_{n\in\mathbb{N}}$ is a subset of A, which is sequentially compact, we can introduce a convergent subsequence by $(n_k)_{k\in\mathbb{N}} \subset \mathbb{N}$, such that

$$\bar{x} = \lim_{k\to\infty} x_{n_k} \in A.$$

From $x_0 \notin A$ it follows that $\bar{x} \neq x_0$, and so there is $e \in E$ such that

$$|\langle x_0 - \bar{x}, e\rangle| \geq 2. \tag{11}$$

On the other hand, since $(e_i)_{i\in\mathbb{N}}$ is dense in E and $x_0 - \bar{x}$ is a continuous linear form, there exists an $i_0 \in \mathbb{N}$ such that

$$\left|\left\langle x_0 - \bar{x}, e - e_{i_0}\right\rangle\right| < 1. \tag{12}$$

Equations (11) and (12) lead to

$$\begin{aligned}\left|\left\langle x_0 - \bar{x}, e_{i_0}\right\rangle\right| &= \left|\langle x_0 - \bar{x}, e\rangle - \left\langle x_0 - \bar{x}, e - e_{i_0}\right\rangle\right| \\ &\geq |\langle x_0 - \bar{x}, e\rangle| - \left|\left\langle x_0 - \bar{x}, e - e_{i_0}\right\rangle\right| > 1.\end{aligned} \tag{13}$$

Combining (13) and (11) gives

$$\begin{aligned}\lim_{k\to\infty}\left|\left\langle x_0 - x_{n_k}, e_{i_0}\right\rangle\right| &= \left|\left\langle x_0 - \bar{x}, e_{i_0}\right\rangle + \lim_{k\to\infty}\left\langle \bar{x} - x_{n_k}, e_{i_0}\right\rangle\right| \\ &= \left|\left\langle x_0 - \bar{x}, e_{i_0}\right\rangle\right| > 1.\end{aligned}$$

Consequently, there is $l_0 \in \mathbb{N}$ such that

$$\left|\left\langle x_0 - x_{n_k}, e_{i_0}\right\rangle\right| > 1, \quad \text{for all } k > l_0. \tag{14}$$

On the other hand, we have assumed that $x_{n_k} \in U_{1n_k}(x_0)$ for all $k \in \mathbb{N}$, i. e.

$$\left|\left\langle x_{n_k} - x_0, e_1\right\rangle\right| < 1, \ \dots\ , \left|\left\langle x_{n_k} - x_0, e_{n_k}\right\rangle\right| < 1.$$

But this contradicts (14), if k is large enough such that $n_k \geq i_0$. ■

Lemma 4 *Let $x_0 \in E'$ and $n \in \mathbb{N}$. Then there is $l \in \mathbb{N}$ such that*

$$x_0 \in U_{2n}\left(e_l'\right), \tag{15}$$

$$U_{2n}(e_l') \subset U_{1n}(x_0). \tag{16}$$

Proof We choose l in a way that

$$e_l' \in U_{2n}(x_0), \tag{17}$$

which means that

$$\left|\left\langle e_l' - x_0, e_k\right\rangle\right| < \tfrac{1}{2} \quad \text{for all } k \in \{1, 2, \dots, n\}.$$

This is possible, since $(e_i')_{i\in\mathbb{N}}$ is dense in E'. Then (15) follows immediately from (17). Now suppose that $y \in U_{2n}(e_l')$, i. e.

$$\left|\left\langle y - e_l', e_k\right\rangle\right| < \tfrac{1}{2} \quad \text{for all } k \in \{1, 2, \dots, n\}.$$

Then we get for all $k \in \{1, 2, \dots, n\}$ that

$$|\langle y - x_0, e_k\rangle| \leq \left|\left\langle y - e_l', e_k\right\rangle\right| + \left|\left\langle e_l' - x_0, e_k\right\rangle\right| < \tfrac{1}{2} + \tfrac{1}{2} = 1,$$

and so $y \in U_{1n}(x_0)$. This implies (16). ■

In order to simplify our notation we define a sequence $(V_i)_{i\in\mathbb{N}}$ of open subsets of E' such that

$$\{V_1, V_2, \dots\} = \left\{U_{2n}\left(e_l'\right) : n, l \in \mathbb{N}\right\}.$$

Combining Lemmas 3 and 4 leads to the following

Corollary 5 *For any sequentially compact set $A \subset E'$ and any $x_0 \in E'$ with $x_0 \notin A$, there exists an index $j \in \mathbb{N}$ such that $x_0 \in V_j$ and $V_j \cap A = \emptyset$.*

Lemma 6 *Assume that both E and E' are separable. Let $X : \Omega \to \mathscr{P}(E')$ be an Effros measurable multifunction with sequentially compact values. Then X is graph measurable, i. e.,* $\mathrm{Graph}\,(X)$ *is $\Sigma \otimes \mathscr{B}(E')$-measurable.*

Proof Recall that the graph of X is defined by

$$\mathrm{Graph}(X) = \left\{(\omega, x) \in \Omega \times E', x \in X(\omega)\right\}.$$

We will show that

$$\mathrm{Graph}\,(X)^{\mathsf{c}} = \bigcup_{n \in \mathbb{N}} X_{-}\,(V_n^{\mathsf{c}}) \times V_n.$$

By Remark 1(a), this implies the measurability of $\mathrm{Graph}(X)$. Assume first that

$$(\omega, x) \in \mathrm{Graph}\,(X)^{\mathsf{c}}.$$

It follows that $(\omega, x) \notin \mathrm{Graph}\,(X)$, and therefore $x \notin X(\omega)$. By assumption, $X(\omega)$ is sequentially compact, and we can apply Corollary 5. Consequently, there is $j \in \mathbb{N}$ such that $V_j \cap X(\omega) = \emptyset$ and $x \in V_j$. Hence, $\omega \notin X^{-}(V_j)$, or equivalently, $\omega \in X_{-}(V_j^{\mathsf{c}})$. We get $(\omega, x) \in X_{-}(V_j^{\mathsf{c}}) \times V_j$ and it follows that

$$(\omega, x) \in \bigcup_{n \in \mathbb{N}} X_{-}(V_n^{\mathsf{c}}) \times V_n. \tag{18}$$

Conversely, assume that (18) holds. Then there is $j \in \mathbb{N}$ such that $(\omega, x) \in X_{-}(V_j^{\mathsf{c}}) \times V_j$. We follow the arguments above in the opposite direction and get $(\omega, x) \in \mathrm{Graph}\,(X)^{\mathsf{c}}$. ■

The final ingredient in the proof of Theorem 2 is the *projection theorem* [4, Theorem III.23].

Projection Theorem *Let (Ω, Σ, P) be a complete probability space, S a Souslin space and H a subset of $\Omega \times S$. If H belongs to $\Sigma \otimes \mathscr{B}(S)$, its projection $\mathrm{pr}_\Omega(H)$ belongs to Σ, where pr_Ω denotes the projection of $\Omega \times S$ onto Ω.*

Proof of Theorem 2. By Lemma 6, the graph of X is measurable, i. e., belongs to $\Sigma \otimes \mathscr{B}(E')$. Let $B \in \mathscr{B}(E')$. Then

$$\mathrm{Graph}\,(X) \cap (\Omega \times B) \in \Sigma \otimes \mathscr{B}(E').$$

Observe that

$$X^{-}(B) = \mathrm{pr}_\Omega \left(\mathrm{Graph}(X) \cap (\Omega \times B)\right).$$

The projection theorem implies that $X^{-}(B) \in \Sigma$. ■

3. Application to the Wave Equation

We address generalized solutions in $\mathscr{D}'(\mathbb{R}^{d+1})$ to the half-space problem for the wave equation

$$\left(\partial_t^2 - c^2\Delta\right) u_c = g, \quad u_c\big|_{t<0} = 0 \tag{19}$$

In other words, the support $\mathrm{supp}\, u_c$ of the solution should be a subset of $\mathbb{R}^d \times [0, \infty)$. Here $g \in \mathscr{D}'(\mathbb{R}^{d+1})$ with $\mathrm{supp}\, g \subset \mathbb{R}^d \times [0, \infty)$ and c is a positive constant.

3.1. The Deterministic Case

It is well known [5, Chapter 6] that the problem (19) has a unique solution $u_c \in \mathscr{D}'(\mathbb{R}^{d+1})$. In particular, for $g = \delta$, the Dirac measure at zero in $\mathbb{R}^{d+1}$, one obtains the unique *fundamental solution* F_c with support in the forward light cone $\{(x,t) \in \mathbb{R}^{d+1} : t \geq 0, |x| \leq ct\}$. If $\mathrm{supp}\, g \subset \mathbb{R}^d \times [0, \infty)$, the convolution of F_c and g exists (Remark 15), and the solution u_c to (19) is given by $u_c = F_c * g$. Explicit formulas for the fundamental solution can be found, e. g., in [22]. In space dimensions $d = 1$ and $d = 2$, F_c is a locally integrable function, in space dimension $d = 3$, it is a Radon measure, and in space dimensions $d \geq 4$, it is a distribution of higher order.

Recall that the Dirac measure in $\mathbb{R}^{d+1}$ is homogeneous of degree -1 with respect to t (Remark 14). With the notation introduced before Remark 14, putting $n = d+1$, this implies that

$$F_c = \tfrac{1}{c} I_c^{-1} F_1$$

or in informal notation

$$F_c(x,t) = \tfrac{1}{c} F_1(x, ct).$$

A rigorous proof of this would involve the tools outlined before Remark 14. Here is an informal proof:

$$\begin{aligned}(\partial_t^2 - c^2\Delta)\tfrac{1}{c} F_1(x,ct) &= \tfrac{1}{c} c^2 \left((\partial_t^2 - \Delta) F_1\right)(x,ct) \\ &= c\delta(x,ct) \;=\; \delta(x,t).\end{aligned}$$

This shows that $\frac{1}{c} I_c^{-1} F_1$ is a fundamental solution of the wave equation with support in the forward light cone. But the fundamental solution (with this property) is unique, hence $\frac{1}{c} I_c^{-1} F_1$ conicides with F_c.

Lemma 7 *The maps $c \to F_c$ and $c \to u_c$ from $(0, \infty)$ to $\mathscr{D}'(\mathbb{R}^{d+1})$ are continuous.*

Proof For the first assertion, observe that the map $(0, \infty) \to \mathscr{D}(\mathbb{R}^{d+1})$, $c \to \frac{1}{c^2} \varphi(x, \frac{t}{c})$ is continuous. Thus

$$c \to \langle \tfrac{1}{c} F_1(x, ct), \varphi(x,t) \rangle = \langle F_1(x,t), \tfrac{1}{c^2} \varphi(x, \tfrac{t}{c}) \rangle$$

is continuous as well. Actually, on can say more: If $c_0 > 0$, the map $[c_0, \infty) \to \mathscr{D}'_\Gamma(\mathbb{R}^{d+1})$, $c \to F_c$, where $\mathscr{D}'_\Gamma(\mathbb{R}^{d+1})$ denotes the space of distributions with support in the cone $\Gamma = \{(x,t) \in \mathbb{R}^{d+1} : t \geq 0, |x| \leq c_0 t\}$, is continuous. To prove the second assertions, it suffices to recall that for fixed $g \in \mathscr{D}'(\mathbb{R}^{d+1})$ with support in the upper half-space, convolution $f \to f * g$ is a continuous map from $\mathscr{D}'_\Gamma(\mathbb{R}^{d+1})$ to $\mathscr{D}'(\mathbb{R}^{d+1})$, see e. g. [24, Section 4] or the arguments in [26, Section 4.1]. ■

3.2. The Stochastic Case

Let (Ω, Σ, P) be a probability space and $d \in \mathbb{N}$. A *generalized stochastic process* X is a weakly measurable map

$$X : \Omega \to \mathscr{D}'(\mathbb{R}^{d+1}),$$

i. e., $\omega \to \langle X(\omega), \varphi \rangle$ is measurable for all $\varphi \in \mathscr{D}(\mathbb{R}^{d+1})$. An important example is *Gaussian space-time white noise*, denoted by $\dot{W}$, with support in $D = \mathbb{R}^d \times [0,\infty)$. Let $\mathscr{S}(D) = \mathscr{S}(\mathbb{R}^{d+1})|D$, where $\mathscr{S}(\mathbb{R}^{d+1})$ is the Schwartz space of rapidly decreasing smooth functions (see Appendix B). Its continuous dual $\Omega = \mathscr{S}'(D)$, equipped with the Borel σ-algebra $\Sigma = \mathscr{B}(\mathscr{S}'(D))$ with respect to the weak topology, is a measurable space. By the *Bochner-Minlos theorem* [7, Section 3.2], [9, Theorem 2.1.1]. there exists a unique probability measure P on $\mathscr{S}'(D)$ such that

$$\int_{\mathscr{S}'(D)} \mathrm{e}^{\mathrm{i}\langle \omega, \varphi \rangle} dP(\omega) = \mathrm{e}^{-\frac{1}{2}\|\varphi\|^2_{L^2(D)}} \tag{20}$$

for all $\varphi \in \mathscr{S}(D)$. The completion of (Ω, Σ, P) is called the *white noise probability space*. Space-time white noise $\dot{W}$ with support in D is the generalized stochastic process

$$\begin{aligned} &\dot{W} : \mathscr{S}'(D) \to \mathscr{D}'(\mathbb{R}^{d+1}), \\ &\langle \dot{W}(\omega), \varphi \rangle = \langle \omega, \varphi|D \rangle, \end{aligned} \tag{21}$$

for all $\varphi \in \mathscr{D}(\mathbb{R}^{d+1})$. By the defining equation (20), $\dot{W}$ is a Gaussian process. Its probabilistic properties are characterized by the Itô isometry

$$\mathrm{E}\langle \dot{W}, \varphi \rangle = 0, \quad \mathrm{E}(\langle \dot{W}, \varphi \rangle)^2 = \|\varphi\|^2_{L^2(D)}.$$

Proposition 8 *There is an (almost surely) unique generalized stochastic process u_c such that*

$$\left(\partial_t^2 - \Delta\right) u_c(\omega) = \dot{W}(\omega), \tag{22}$$

$$\operatorname{supp} u(\omega) \subset \mathbb{R}^d \times [0,\infty), \tag{23}$$

in $\mathscr{D}'\left(\mathbb{R}^{d+1}\right)$ for (almost) all $\omega \in \Omega$. Further $u_c(\omega)$ depends continuously on $c \in (0,\infty)$.

Proof Pathwise existence and uniqueness follows from the deterministic theory outlined in Subsection 3.1. The solution is given by

$$u_c(\omega) = F_c * \dot{W}(\omega)$$

and depends continuously on c by Lemma 7. Measurability at fixed c follows from the fact that u_c is obtained as a composition of a measurable and a continuous map:

$$\omega \to \dot{W}(\omega) \to F_c * \dot{W}(\omega),$$

see again Lemma 7. ■

Modelling the uncertainty in the propagation speed by a compact interval $[\underline{c}, \overline{c}]$ with $\underline{c} > 0$, we are now in the position to define the solution to the stochastic wave equation as a multifunction

$$\begin{aligned} &X : \Omega \to \mathscr{P}\left(\mathscr{D}'(\mathbb{R}^{d+1})\right) \\ &X(\omega) = \{u_c(\omega) : \underline{c} \le c \le \overline{c}\}. \end{aligned} \tag{24}$$

As continuous images of a compact interval, the values $X(\omega)$ are sequentially compact and, in particular, closed.

Lemma 9 *The multifunction (24) admits a Castaing representation and is Effros measurable.*

Proof Let $(c_i)_{i\in\mathbb{N}}$ be a sequence such that $\mathrm{cl}\{c_i : i \in \mathbb{N}\} = [\underline{c}, \overline{c}]$. Then, because of the continuity of the function $c \to u_c(\omega)$,

$$X(\omega) = \mathrm{cl}\left\{u_{c_i}(\omega) : i \in \mathbb{N}\right\}$$

in $\mathscr{D}'(\mathbb{R}^{d+1})$, for (almost) all ω. By the measurability of the u_{c_i} we conclude that $(u_{c_i})_{i\in\mathbb{N}}$ is a Castaing representation of X. By Remark 1(b), X is Effros measurable. ■

Here is the main result of this section.

Theorem 10 *The multifunction (24) is a random set.*

Proof We wish to apply Theorem 2 with $E = \mathscr{D}(\mathbb{R}^{n+1})$, $E' = \mathscr{D}'(\mathbb{R}^{n+1})$. By [23, Appendix], both $\mathscr{D}(\mathbb{R}^{n+1})$ and $\mathscr{D}'(\mathbb{R}^{n+1})$ are Souslin spaces and, in particular, separable. The multifunction X is Effros measurable by Lemma 9 and has sequentially compact values. The hypotheses of Theorem 2 are satisfied. ■

4. The One-Dimensional Wave Equation

In one space dimension, the fundamental solution is given by

$$F_c(x,t) = \tfrac{1}{2c} H(t) H(ct - |x|)$$

where H denotes the Heaviside function. The convolution of F_c with a function g with support in the upper half-plane results in d'Alembert's formula

$$u_c(x,t) = \frac{1}{2c} \int_0^t \int_{x-cs}^{x+cs} g(y,s)\, dy\, ds$$

for the solution of the wave equation (19). This formula can be rewritten as follows. Denote by $\chi_{(x,t,c)}$ the indicator function of the backward triangle $\{(y,s) : 0 \le s \le t, x - cs \le y \le x + cs\}$ with vertex at (x,t). Then $u_c(x,t) = \langle g, \frac{1}{2c}\chi_{(x,t,c)} \rangle$.

4.1. Random Field Solution in One Dimension

The solution to the one-dimensional ($d = 1$) stochastic wave equation (22), (23) is known [25] to be a scaled Brownian sheet, transformed to the axis $x = \pm ct$. Actually, by the Itô isometry, the action of $\dot{W}$ on test functions can be extended to functions in $L^2(\mathbb{R}^2)$. Thus the solution to the one-dimensional stochastic wave equation can be written as

$$u_c(x,t) = \langle \dot{W}, \tfrac{1}{2c}\chi_{(x,t,c)}\rangle = \tfrac{1}{2c}\int_0^t \int_{x-cs}^{x+cs} dW(y,s).$$

Lemma 11 *Let $(x,t) \in \mathbb{R} \times (0,\infty)$, $c_1, c_2 > 0$. Then the covariance between $u_{c_1}(x,t)$ and $u_{c_2}(x,t)$ is given by*

$$\mathrm{E}(u_{c_1}(x,t)u_{c_2}(x,t)) = \frac{t^2}{4}\min\left(\frac{1}{c_1}, \frac{1}{c_2}\right).$$

Proof Again by the Itô isometry,

$$\begin{aligned}&\mathrm{E}(u_{c_1}(x,t)u_{c_2}(x,t))\\ &= \tfrac{1}{4c_1c_2}\int_0^\infty \int_{-\infty}^\infty \chi_{(x,t,c_1)}(y,s)\chi_{(x,t,c_2)}(y,s)dyds.\end{aligned}$$

The latter integral equals the area of the smaller triangle, thus it is $t^2 \min(c_1, c_2)$. After division by $4c_1c_2$, the smaller constant cancels, and this results in the desired formula. ■

In particular, $u_c(x,t)$ vanishes as $c \to \infty$ in the mean square sense, i. e.,

$$\lim_{c\to\infty} \mathrm{E}(u_c^2(x,t)) = 0.$$

Proposition 12 *Let $t > 0$ and $x \in \mathbb{R}$ be fixed. The stochastic process $(v_r)_{r\geq 0}$, defined by*

$$v_r = \frac{2}{t}u_{\frac{1}{r}}(t,x), \quad r > 0 \tag{25}$$

with $v_0 = 0$ is a Brownian motion.

Proof Clearly v is a Gaussian process with zero expectation. The covariance we get from Lemma 11 as

$$\mathrm{E}(v_r v_s) = \frac{4}{t^2}\mathrm{E}\left(u_{\frac{1}{r}}(t,x)u_{\frac{1}{s}}(t,x)\right) = \min(r,s).$$

These properties characterize Brownian motion. ■

Proposition 13 *Let $\underline{c}$ and $\overline{c}$ be two real numbers such that*

$$0 < \underline{c} < \overline{c} < \infty.$$

Then, for fixed $t > 0$ and $x \in \mathbb{R}$, the multifunction

$$\begin{aligned}X : \Omega &\to \mathscr{P}(\mathbb{R}),\\ \omega &\to \{u_c(x,t,\omega) : \underline{c} \leq c \leq \overline{c}\},\end{aligned}$$

is a random set.

Proof From (25) we get

$$u_c(t,x) = \frac{t}{2}v_{\frac{1}{c}}$$

almost surely. Since v is a Brownian motion, the function $c \to u_c(x,t,\omega)$ is continuous for almost all ω. It follows that the sequence of random variables $(u_{c_n}(t,x))_{n\in\mathbb{N}}$ is a Castaing representation of X, where $(c_n)_{n\in\mathbb{N}}$ is a sequence such that $\mathrm{cl}\{c_n : n \in \mathbb{N}\} = [\underline{c}, \overline{c}]$. Applying the Fundamental Measurability Theorem in its classical form for Polish spaces (mentioned in Section 2 after Remark 1) implies that X is Borel measurable and therefore a random set. ■

4.2. Upper and Lower Probabilities

In this section, we are going to compute the upper probability $\overline{P}(B)$ and lower probability $\underline{P}(B)$ for any open interval $B = (\underline{b}, \overline{b})$. By definition, the upper probability equals

$$\overline{P}(B) = P(X \cap B \neq \emptyset). \tag{26}$$

Applying the theorem of total probability we can write

$$\begin{aligned}&P(X \cap B \neq \emptyset) = P\left(\exists c \in [\underline{c}, \overline{c}] : \underline{b} < u_c(t,x) < \overline{b}\right)\\ &= \int_{-\infty}^{\underline{b}} P\left(\exists c \in [\underline{c}, \overline{c}] : u_c(t,x) > \underline{b} \middle| u_{\overline{c}}(t,x) = y\right) f(y)dy\\ &+ \int_{\underline{b}}^{\overline{b}} f(y)dy\\ &+ \int_{\overline{b}}^{\infty} P\left(\exists c \in [\underline{c}, \overline{c}] : u_c(t,x) < \overline{b} \middle| u_{\overline{c}}(t,x) = y\right) f(y)dy,\end{aligned}$$

where f is the probability density of the random variable $u_{\overline{c}}(t,x)$. Letting $\underline{r} = 1/\overline{c}$, $\overline{r} = 1/\underline{c}$, we get

$$\begin{aligned}&P(X \cap B \neq \emptyset)\\ &= \int_{-\infty}^{\frac{2}{t}\underline{b}} P\left(\exists r \in [\underline{r}, \overline{r}] : v_r > \frac{2}{t}\underline{b} \middle| v_{\underline{r}} = y\right) g(y)dy\\ &+ \int_{\frac{2}{t}\underline{b}}^{\frac{2}{t}\overline{b}} g(y)dy\\ &+ \int_{\frac{2}{t}\overline{b}}^{\infty} P\left(\exists r \in [\underline{r}, \overline{r}] : v_r < \frac{2}{t}\overline{b} \middle| v_{\underline{r}} = y\right) g(y)dy,\end{aligned} \tag{27}$$

where

$$g(y) = \frac{1}{\sqrt{2\pi\underline{r}}}e^{-\frac{y^2}{2\underline{r}}}$$

is the probability density of the random variable $v_{\underline{r}}(t,x)$. It is a Gaussian random variable with variance $\underline{r}$. The probabilities in (27) we express by first hitting times of a Brownian motion.

So let $(w_t)_{t\in[0,\infty)}$ be a standard Brownian motion starting at 0 and let $a \in \mathbb{R}$. The first hitting time $\tau(a)$ is defined by

$$\tau(a) := \min\{t : w_t = a\}.$$

Its probability distribution $F_{\tau(a)}$ is well known [2, Section 7.4]:

$$F_{\tau(a)}(t) = P(\tau(a) \le t) = \frac{2}{\sqrt{2\pi t}} \int_{|a|}^{\infty} e^{-\frac{u^2}{2t}}\, du.$$

If $a > 0$, then

$$P(\exists s \in [0,t] : w_s \ge a) = F_{\tau(a)}(t).$$

From the continuity of $F_{\tau(a)}(t)$ with respect to a it follows that

$$\begin{aligned} P(\exists s \in [0,t] : w_s > a) &= \lim_{\varepsilon \to 0} P(\exists s \in [0,t] : w_s \ge a + |\varepsilon|) \\ &= \lim_{\varepsilon \to 0} F_{\tau(a+|\varepsilon|)}(t) = F_{\tau(a)}(t). \end{aligned}$$

Since v_r is a Brownian motion with respect to r, it follows that

$$\begin{aligned} &P\left(\exists r \in [\underline{r}, \overline{r}] : v_r > \frac{2}{t}\underline{b} \middle| v_{\underline{r}} = y\right) \\ &= F_{\tau\left(\frac{2}{t}\underline{b} - y\right)}(\overline{r} - \underline{r}) \end{aligned}$$

and

$$\begin{aligned} &P\left(\exists r \in [\underline{r}, \overline{r}] : v_r < \frac{2}{t}\overline{b} \middle| v_{\underline{r}} = y\right) \\ &= F_{\tau\left(\frac{2}{t}\overline{b} - y\right)}(\overline{r} - \underline{r}). \end{aligned}$$

Hence we can rewrite (27) as

$$\begin{aligned} P(X \cap B) \ne \emptyset) = &\int_{-\infty}^{\frac{2}{t}\underline{b}} F_{\tau\left(\frac{2}{t}\underline{b} - y\right)}(\overline{r} - \underline{r})\, g(y) dy \\ &+ \int_{\frac{2}{t}\underline{b}}^{\frac{2}{t}\overline{b}} g(y) dy \\ &+ \int_{\frac{2}{t}\overline{b}}^{\infty} F_{\tau\left(\frac{2}{t}\overline{b} - y\right)}(\overline{r} - \underline{r})\, g(y) dy. \end{aligned}$$

Inserting this into (26) shows that the upper probability $\overline{P}(B)$ can be expressed in terms of first hitting times of a Brownian motion.

Finally, we derive a similar expression for the lower probability

$$\underline{P}(B) = P(X \subset B).$$

Using again the law of total probability gives that

$$\underline{P}(B) = P\left(\tfrac{2}{t}\underline{b} < v_r < \tfrac{2}{t}\overline{b}, \forall r \in [\underline{r}, \overline{r}]\right)$$

equals

$$\int_{\frac{2}{t}\underline{b}}^{\frac{2}{t}\overline{b}} P\left(\tfrac{2}{t}\underline{b} < v_r < \frac{2}{t}\overline{b}, \forall r \in [\underline{r}, \overline{r}] \middle| v_{\underline{r}} = y\right) g(y) dy. \tag{28}$$

The first exit time $\tau(a,b)$ of a Brownian motion $(w_t)_{t \in (0,\infty)}$ for $a < 0$ and $b > 0$ is defined by

$$\tau(a,b) := \min\{t : w_t \notin (a,b)\}.$$

It has the probability distribution

$$F_{\tau(a,b)}(t) = \int_0^t \mathrm{cc}_s\left(\frac{b+a}{2}, \frac{b-a}{2}\right) ds,$$

where

$$\mathrm{cc}_s(x,y) = \mathscr{L}^{-1}_{\gamma \to s}\left(\frac{\cosh(x\sqrt{2\gamma})}{\cosh(y\sqrt{2\gamma})}\right)$$

for $x < y$ [3, p. 212 and 641]. Here $\mathscr{L}^{-1}$ denotes the inverse Laplace transform. By the same arguments as in the case of $\overline{P}$ we can rewrite (28) as

$$\underline{P}(B) = \int_{\frac{2}{t}\underline{b}}^{\frac{2}{t}\overline{b}} \left(1 - F_{\tau\left(\frac{2}{t}\underline{b} - y, \frac{2}{t}\overline{b} - y\right)}(\overline{r} - \underline{r})\right) g(y) dy.$$

Conclusion

The solution to the stochastic wave equation with space-time white noise as excitation is a generalized stochastic process when the space dimension d is greater or equal to 2. Set-valued solutions are obtained when the propagation speed is modelled as an interval. In order to establish that the set-valued solution is a random set, a new measurability theorem in non-metrizable spaces was required. In one space dimension, the solution to the stochastic wave equation is a scaled and rotated Brownian sheet. In addition, the map $c \to u_c(x,t)$ is a classical stochastic process; the covariances $\mathrm{E}(u_{c_1}(x,t), u_{c_2}(x,t))$ can be computed classically, and the process $c \to u_c(x,t)$ can be transformed into a Brownian motion, from where upper and lower probabilities could be computed.

In higher space dimensions, one still can compute the covariances of the processes $c \to \langle u_c, \varphi \rangle$ for $\varphi \in \mathscr{D}(\mathbb{R}^{d+1})$. In dimensions $d = 2$ and $d = 3$ this was done in [26]. This information might be used to compute upper and lower probabilities of certain functionals of the solution. An example of such a functional could be the integral of u_c over a compact subset of $\mathbb{R}^{d+1}$, which can be defined by employing the Itô isometry. This is a topic for future research.

Appendix A. Topological Vector Spaces

For the convenience of the reader, the required notions from the theory of topological vector spaces and the theory of distributions are collected in these two appendices. All details on the topics of the two appendices can be found, for example, in [10, 23].

A vector space E (over $\mathbb{R}$ or $\mathbb{C}$) is a *topological vector space*, if it is equipped with a topology making addition and scalar multiplication continuous. It is a *Hausdorff space* if any two points have non-intersecting neighborhoods.

The *dual* E' of E is the vector space of all continuous linear maps from E into $\mathbb{R}$ (or $\mathbb{C}$). The action of a linear map $x \in E'$ on an element $e \in E$ is usually denoted by the *duality bracket*

$$x(e) = \langle x, e\rangle.$$

The *weak topology* on E' is given as follows. For $x \in E'$, a base of neighborhoods is given by the sets

$$\{y \in E' : |\langle x-y, e_1\rangle| < \varepsilon, \ldots, |\langle x-y, e_n\rangle| < \varepsilon\}$$

where $n \in \mathbb{N}$, $e_1, \ldots, e_n \in E$, and $\varepsilon > 0$. In the definition, ε may be taken equal to 1 without loss of generality upon replacing e_i by e_i/ε.

A topological vector space E is *locally convex* if every point in E has a base of convex neighborhoods. Locally convex topological vector spaces always have rich duals. (This is not so for non-locally convex topological vector spaces: examples are known where the dual degenerates to $E' = \{0\}$.) While Theorem 2 holds for general topological vector spaces, local convexity is an important ingredient in the theory of distributions.

Appendix B. Schwartz Distributions

The theory of generalized functions of Laurent Schwartz, the *theory of distributions*, has become *the* standard in the theory of linear partial differential equations. It is needed here in order to give a meaning to generalized stochastic processes such as space-time white noise in the context of the wave equation. Generally, distributions are defined as members of the dual of certain spaces of *test functions*. Let O be an open subset of $\mathbb{R}^n$. The vector space $\mathscr{D}(O)$ is the space of infinitely differentiable (real or complex valued) functions on O of compact support, i.e., vanishing outside some compact subset of O. Given a compact subset $K \subset \Omega$, $\mathscr{D}_K(O)$ denotes the subspace of infinitely differentiable functions with support in K. Equipped with the topology of uniform convergence, together with all derivatives, $\mathscr{D}_K(O)$ becomes a metrizable, complete locally convex space. Clearly, $\mathscr{D}(O)$ is the union of the spaces $\mathscr{D}_K(O)$ as K runs through the compact subsets of O. It is equipped with the locally convex inductive limit topology, that is, the finest locally convex topology that makes all inclusions $\mathscr{D}_K(O) \to \mathscr{D}(O)$ continuous.

The *space of distributions* $\mathscr{D}'(O)$ is the dual space of $\mathscr{D}(O)$. Usually, distributions are denoted by roman letters, whereas test functions are denoted by greek letters. The action of a distribution u on a test function φ is denoted by the duality bracket. A common abuse of notation is to display the variable $x \in O$ in the duality bracket. Thus the following notations are in use:

$$u(\varphi) = \langle u, \varphi\rangle = \langle u(x), \varphi(x)\rangle$$

where the last equality is formally incorrect, albeit quite intuitive in certain situations.

Every locally integrable function f and every Radon measure μ on O can be viewed as a distribution by means of the prescriptions

$$\langle f, \varphi\rangle = \int_O \varphi(x)dx, \quad \text{respectively,} \quad \int_O \varphi(x)d\mu(x).$$

In particular, the *Dirac measure at the origin* acts as $\langle \delta, \varphi\rangle = \varphi(0)$.

Distributions can be restricted to open subsets. If $O_1 \subset O$ is open and $u \in \mathscr{D}'(O)$, the *restriction* is defined by $\langle u|O_1, \varphi\rangle = \langle u, \varphi\rangle$ where $\varphi \in \mathscr{D}(O_1)$, viewed as a subspace of $\mathscr{D}(O)$. The *support* of a distribution $u \in \mathscr{D}'(O)$, denoted by $\operatorname{supp} u$, is the complement of the largest open subset of O restricted to which u is equal to 0.

Derivatives. The *partial derivatives* of a distribution $u \in \mathscr{D}'(O)$, sometimes referred to as *weak derivatives* are defined by

$$\langle \partial_{x_i} u, \varphi\rangle = -\langle u, \partial_{x_i}\varphi\rangle, \quad i = 1, \ldots, n.$$

Higher order derivatives are defined recursively.

Multiplication by smooth functions. If $u \in \mathscr{D}'(O)$ and χ is an infinitely differentiable function on O, the product $\chi u \in \mathscr{D}'(O)$ is defined by

$$\langle \chi u, \varphi\rangle = \langle u, \chi\varphi\rangle.$$

Let $\tau_h : x \to x + h$ be the translation of $\mathbb{R}^n$ by h and A an invertible $(n \times n)$-matrix with the corresponding linear map $x \to Ax$ from $\mathbb{R}^n$ to $\mathbb{R}^n$. These operations act on $\mathscr{D}(O)$ by

$$\tau_h\varphi(x) = \varphi(x-h), \quad A\varphi(x) = \varphi(A^{-1}x)$$

and on $\mathscr{D}'(O)$ by

$$\langle \tau_h u, \varphi\rangle = \langle u, \tau_{-h}\varphi\rangle, \quad \langle Au, \varphi\rangle = \langle u, |\det A|A^{-1}\varphi\rangle.$$

These formulas are much easier to remember, employing the informal notation introduced above:

$$\langle \tau_h u, \varphi\rangle = \langle u(x-h), \varphi(x)\rangle = \langle u(x), \varphi(x+h)\rangle$$
$$\langle Au, \varphi\rangle = \langle u(A^{-1}x), \varphi(x)\rangle = \langle u(x), |\det A|\varphi(Ax)\rangle.$$

Let $c \neq 0$ and I_c the $(n \times n)$-identity matrix, where the last diagonal entry is replaced by c. Then I_c maps $\mathbb{R}^n$ into $\mathbb{R}^n$, or stated expicitly, $I_c(x_1, x_2, \ldots, x_n) = (x_1, \ldots, x_{n-1}, cx_n)$. A distribution on $u \in \mathscr{D}'(\mathbb{R}^n)$ is *homogeneous of degree* $k \in \mathbb{R}$ *with respect to the last variable*, if $I_c^{-1}u = c^k u$ for all $c \neq 0$. Informally written, this means that $u(x_1, \ldots, x_{n-1}, cx_n) = c^k u(x_1, \ldots, x_{n-1}, x_n)$.

Remark 14 *The Dirac measure on $\mathbb{R}^n$ is homogeneous of degree -1 with respect to the last (or any other) variable. Indeed,*

$$\langle I_c^{-1}\delta, \varphi\rangle = \langle \delta, \tfrac{1}{c}I_c\varphi\rangle = \tfrac{1}{c}\langle \delta, \varphi\rangle$$

In informal notation, the derivation looks as follows:

$$\begin{aligned}&\langle \delta(x_1, \ldots, x_{n-1}, cx_n), \varphi(x_1, \ldots, x_{n-1}, x_n)\rangle \\ &= \langle \delta(x_1, \ldots, x_{n-1}, x_n), \tfrac{1}{c}\varphi(x_1, \ldots, x_{n-1}, x_n/c)\rangle \\ &= \tfrac{1}{c}\varphi(0, \ldots, 0, 0)\rangle = \tfrac{1}{c}\langle \delta, \varphi\rangle.\end{aligned}$$

Convolution. The convolution of a distribution $u \in \mathscr{D}'(\mathbb{R}^n)$ with a test function $\varphi \in \mathscr{D}(\mathbb{R}^n)$ at point $x \in \mathbb{R}^n$ is defined by

$$(u * \varphi)(x) = \langle u, \tau_{-x}\varphi \rangle$$

or in informal notation,

$$(u * \varphi)(x) = \langle u(y), \varphi(x+y) \rangle.$$

It is an infinitely differentiable function of $x \in \mathbb{R}^n$. The convolution of a distribution $u \in \mathscr{D}'(\mathbb{R}^n)$ and a compactly supported distribution $v \in \mathscr{D}'(\mathbb{R}^n)$ is defined by

$$\langle u * v, \varphi \rangle = \langle u, \check{v} * \varphi \rangle$$

where $\check{v}$ is defined by $\check{v}(y) = v(-y)$. In order to show that this definition makes sense, one has to invoke the fact that the support of $\check{v} * \varphi$ is contained in the (Minkowski) sum of the supports of φ and $\check{v}$. Thus $\check{v} * \varphi$ belongs to $\mathscr{D}(\mathbb{R}^n)$ and may indeed serve as an argument of the distribution u.

Remark 15 *This definition can be generalized to* distributions whose supports are in favorable position*, meaning that the intersection of the supports of u and $\check{v} * \varphi$ are compact for every $\varphi \in \mathscr{D}(O)$. In this case, the convolution is defined by*

$$\langle u * v, \varphi \rangle = \langle u, \chi(\check{v} * \varphi) \rangle$$

where χ is an arbitrary element of $\mathscr{D}(\mathbb{R}^n)$ which is identically equal to 1 in a neighborhood of $\operatorname{supp} u \cap \operatorname{supp}(\check{v} * \varphi)$*. This situation arises, for example, when u has its support in a closed half space D and v has its support in a closed acute cone Γ contained in the interior of D.*

Tempered distributions. An infinitely differentiable function ψ on $\mathbb{R}^n$ is called *rapidly decreasing*, if all its derivatives decay faster to zero than any negative power of $|x|$ as $|x| \to \infty$. The space of rapidly decreasing functions $\mathscr{S}(\mathbb{R}^n)$ is topologized by the (countable) family of seminorms given by the suprema over $\mathbb{R}^n$ of the derivatives of ψ, weighted by $(1+|x|)^m$, $m \in \mathbb{N}$. Its dual $\mathscr{S}'(\mathbb{R}^n)$ is the space of *tempered distributions*. It is a subspace of $\mathscr{D}'(\mathbb{R}^n)$. If $D \subset \mathbb{R}^n$ is a half space (or – as a matter of fact – the closure of an open subset of $\mathbb{R}^n$ with a sufficiently regular boundary) one may define $\mathscr{S}(D)$ as the set of restrictions of elements of $\mathscr{S}(\mathbb{R}^n)$, with the induced topology. Its dual $\mathscr{S}'(D)$ can be identified with the space of tempered distributions on $\mathbb{R}^n$ with support in D.

To complete this excursion into distribution spaces, we are going to sketch a proof why $\mathscr{D}'(O)$ is not metrizable with respect to the weak topology. We will have to invoke a few further results from the theory of topological vector spaces. The basic observation is that $\mathscr{D}'(O)$ is sequentially complete, that is, every weak Cauchy sequence converges. This follows from [10, Corollary to Proposition 3.6.5], together with [10, Example 3.6.4]. Thus, if $\mathscr{D}'(O)$ were metrizable, it would be a complete locally convex topological vector space (because on metrizable spaces, completeness and sequential completeness coincide). From there, contradictions can be obtained in various ways. A short route is to infer that in this case $\mathscr{D}'(O)$ would be equal to the algebraic dual $\mathscr{D}^*(O)$, which is not true. Indeed, let E be a vector space (over $\mathbb{R}$ or $\mathbb{C}$). The algebraic dual E^* is the vector space of all linear maps from E into $\mathbb{R}$ (or $\mathbb{C}$). It is known that the continuous dual E' of E is complete with the weak topology if and only if it coincides with the algebraic dual, that is, $E' = E^*$, see [17, Chapter IV, Section 6]. Applied in our situation, this would mean that every linear form on $\mathscr{D}(O)$ is continuous. However, this is not the case. For example, one may define the linear form $\langle u, \varphi \rangle = \sum_{k=0}^{\infty} |\partial^k \varphi|$ first on the subspace of all $\varphi \in \mathscr{D}(\mathbb{R})$ for which this sum is finite and then extend u algebraically to all of $\mathscr{D}(\mathbb{R})$. Involving infinitely many derivatives at zero, u does not belong to $\mathscr{D}'(\mathbb{R})$, as follows from [10, Proposition 4.4.5].

Acknowledgments

This work was supported by the FWF doctoral program Computational Interdisciplinary Modelling W1227 as well as the research project P-27570-N26 of FWF (The Austrian Science Fund).

References

[1] Diego Averna. Separation properties in X and 2^X. Upper semicontinuous and measurable multifunctions. *Rendiconti del Circolo Matematico di Palermo. Serie II*, 38(1):140–151, 1989.

[2] Frank Beichelt. *Stochastic Processes in Science, Engineering and Finance*. Chapman & Hall/CRC, Boca Raton, FL, 2006.

[3] Andrei N. Borodin and Paavo Salminen. *Handbook of Brownian motion—facts and formulae*. Probability and its Applications. Birkhäuser Verlag, Basel, second edition, 2002.

[4] Charles Castaing and Michel Valadier. *Convex analysis and measurable multifunctions*, volume 580 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin-New York, 1977.

[5] Jaque Chazarain and Alain Piriou. *Introduction to the Theory of Linear Partial Differential Equations*. North-Holland, Amsterdam, 1982.

[6] Robert Dalang, Davar Khoshnevisan, Carl Mueller, David Nualart, and Yimin Xiao. *A minicourse on stochastic partial differential equations*, volume 1962 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 2009.

[7] Takeyuki Hida. *Brownian motion*, volume 11 of *Applications of Mathematics*. Springer-Verlag, New York-Berlin, 1980. Translated from the Japanese by the author and T. P. Speed.

[8] Charles J. Himmelberg. Measurable relations. *Polska Akademia Nauk. Fundamenta Mathematicae*, 87:53–72, 1975.

[9] Helge Holden, Bernt Øksendal, Jan Ubøe, and Tusheng Zhang. *Stochastic partial differential equations*. Universitext. Springer, New York, second edition, 2010. A modeling, white noise functional approach.

[10] John Horváth. *Topological vector spaces and distributions. Vol. I*. Addison-Wesley Publishing Co., Reading, Massachusetts, 1966.

[11] Jun Gang Li, Shou Mei Li, and Yukio Ogura. Strong solution of Itô type set-valued stochastic differential equation. *Acta Mathematica Sinica-English Series*, 26(9):1739–1748, 2010.

[12] Thomas Lorenz. *Mutational analysis*, volume 1996 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 2010.

[13] Marek T. Malinowski and Mariusz Michta. The interrelation between stochastic differential inclusions and set-valued stochastic differential equations. *Journal of Mathematical Analysis and Applications*, 408(2): 733–743, 2013.

[14] Ilya Molchanov. *Theory of random sets*. Probability and its Applications (New York). Springer-Verlag London, Ltd., London, 2005.

[15] Michael Oberguggenberger. Engineering. In Thomas Augustin, Frank P. A. Coolen, Gert de Cooman, and Matthias C. M. Troffaes, editors, *Introduction to imprecise probabilities*, pages 291–304. John Wiley & Sons, Ltd., Chichester, 2014.

[16] Michael Oberguggenberger. Analysis and computation with hybrid random set stochastic models. *Structural Safety*, 52:233–243, 2015.

[17] Helmut H. Schaefer. *Topological vector spaces*. Springer-Verlag, New York-Berlin, 1980. Fourth printing corrected.

[18] Bernhard Schmelzer. On solutions of stochastic differential equations with parameters modeled by random sets. *International Journal of Approximate Reasoning*, 51(9):1159–1171, 2010.

[19] Bernhard Schmelzer. Set-valued assessments of solutions to stochastic differential equations with random set parameters. *Journal of Mathematical Analysis and Applications*, 400(2):425–438, 2013.

[20] Bernhard Schmelzer, Michael Oberguggenberger, and Christoph Adam. Efficiency of tuned mass dampers with uncertain parameters on the performance of structures under stochastic excitation. *Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability*, 224(O4):297–308, 2010.

[21] Björn Sprungk and K. Gerald van den Boogaart. Stochastic differential equations with fuzzy drift and diffusion. *Fuzzy Sets and Systems*, 230:53–64, 2013.

[22] François Trèves. *Basic linear partial differential equations*, volume 62 of *Pure and Applied Mathematics*. Academic Press, New York-London, 1975.

[23] François Trèves. *Topological vector spaces, distributions and kernels*. Academic Press, New York-London, 1967.

[24] Vasilij S. Vladimirov. *Distributions en physique mathématique*. Édition Mir, Moscow, 1979.

[25] John B. Walsh. An introduction to stochastic partial differential equations. In Paul Louis Hennequin, editor, *École d'été de probabilités de Saint-Flour, XIV—1984*, volume 1180 of *Lecture Notes in Mathematics*, pages 265–439. Springer, Berlin, 1986.

[26] Lukas Wurzer. *Stochastic Differential Equations, Random Sets and Distributions*. Universität Innsbruck, 2015. PhD dissertation.