7. White noise and the Paley-Wiener integral

Another important source of motivation for the introduction of the stochastic integral comes from modeling of noise in physical phenomena. Recall that a classical way to describe the evolution of a quantity X is by relating its infinitesimal rate of change to its value. This represents the evolution by the ODE

$$\frac{\mathrm{d}X_t}{\mathrm{d}t} = a(t, X_t) \tag{7.1}$$

where $a\colon [0,\infty)\times\mathbb{R}\to\mathbb{R}$ is determined by the specific context. This equation is idealized because in realistic situations, another term appears on the right-hand side reflecting unpredictable, or outright random, influences of ambient environment, thus upgrading (7.1) to the form

$$\frac{\mathrm{d}X_t}{\mathrm{d}t} = a(t, X_t) + W_t \tag{7.2}$$

where $\widetilde{W} = \{W_t\colon t\in[0,\infty)\}$ is a "noise" process.

The exact nature of W is determined by the specific context, but the simplest requirements we could impose are that W be

(1) of zero mean, $EW_t = 0$, to ensure that all deterministic effects are already accounted by the first term on the right-hand side of (7.2), and
(2) independent at different times, $\forall t\neq s\colon W_t \perp\!\!\!\perp W_s$, or at least uncorrelated $\forall t\neq s\colon \operatorname{Cov}(W_s, W_t) = 0$, to express that the noise is memoryless.

In order to solve the ODE, we then also need require that $t\mapsto W_t$ is at least measurable and therein lies a problem: No process satisfying these requirements exists. (More precisely, while the Kolmogorov Extension Theorem readily outputs a process satisfying (1-2) for any prescribed distribution for each W_t, as soon as these distributions are non-degenerate, the sample paths of this process will fail to be measurable.)

7.1 Poisson point process.

There are several ways how to overcome this problem. One of these is to give up on independence at different times and allow for non-trivial correlations. The other is to interpret (7.2) only under an integral sign. Instead of a value attached to a point, W then just needs to assign a value to each non-degenerate interval, or more generally, to each measurable set of times and that so in an additive way. In short, W becomes a signed measure.

One example of such an object is the *compensated Poisson process*. Recall the following definition from 275B:

Definition 7.1 (Poisson point process) *Let $(\mathscr{X}, \Sigma, \mu)$ be a measure space with μ finite. A Poisson point process on $\mathscr{X}$ with intensity measure μ is a collection $\{N(A)\colon A\in\Sigma\}$ of random variables such that*

$$A\mapsto N(A) \text{ is countably additive on } \Sigma \tag{7.3}$$

and the following is true:

$$\forall n \geqslant 1\, \forall A_1,\dots,A_n \in \Sigma \text{ disjoint}: \quad \{N(A_i)\}_{i=1}^n \text{ are independent} \tag{7.4}$$

and

$$\forall A \in \Sigma: \quad N(A) = \text{Poisson}(\mu(A)) \tag{7.5}$$

In class we use a slightly weaker definition in which (7.3) is replaced by

$$\forall \{A_i\}_{i\in\mathbb{N}} \in \Sigma: \text{ disjoint } \Rightarrow N\Bigl(\bigcup_{i\in\mathbb{N}} A_i\Bigr) = \sum_{i\in\mathbb{N}} N(A_i) \quad \text{a.s.} \tag{7.6}$$

where the implicit null event is allowed to depend on $\{A_i\}_{i\in\mathbb{N}}$. That weakening (7.3) carries no gain or loss is seen from:

Lemma 7.2 *Let K be Poisson with parameter $\mu(\mathscr{X})$ and let $\{X_i\}_{i\in\mathbb{N}}$ be i.i.d. $\mathscr{X}$-valued with law $\mu(\cdot)/\mu(\mathscr{X})$, independent of K. Then $\{N(A)\colon A \in \Sigma\}$ with*

$$N(A) := \sum_{i=1}^{K} 1_A(X_i) \tag{7.7}$$

is a Poisson point process with intensity μ according to Definition 7.1.

Proof (sketch). Obviously, $A \mapsto N(A)$ is countably additive (and thus a random measure, in a suitable interpretation of that term). We thus have to check that (7.4) and (7.5) hold. Let $A \in \Sigma$. Using that $N(A)$ is integer valued, given $k \in \mathbb{N}$, we compute

$$\begin{aligned} P\bigl(N(A)=k\bigr) &= \sum_{n=0}^{\infty} P\bigl(K=n\bigr) P\Biggl(\sum_{i=1}^{n} 1_A(X_i) = k\Biggr) \\ &= \sum_{n=k}^{\infty} \frac{\mu(\mathscr{X})^n}{n!} \mathrm{e}^{-\mu(\mathscr{X})} \binom{n}{k} \Bigl(\frac{\mu(A)}{\mu(\mathscr{X})}\Bigr)^k \Bigl(\frac{\mu(A^{\mathrm{c}})}{\mu(\mathscr{X})}\Bigr)^{n-k} \end{aligned} \tag{7.8}$$

Here we used that K and $\{X_i\}_{i\in\mathbb{N}}$ are independent and then, in the second line, wrote up the explicit probability mass function of K and use the fact that, due to $\{X_i\}_{i\in\mathbb{N}}$ being i.i.d., $\{1_A(X_i)\}_{i=1}^n$ are i.i.d. with parameter $\mu(A)/\mu(\mathscr{X})$. The resulting expression is now readily shown to equal $\frac{1}{k!}\mu(A)^k \mathrm{e}^{-\mu(A)}$ giving us (7.5). The proof of (7.4) is analogous, albeit notationally more involved. We leave the details to the reader. □

Remark 7.3 The above construction allows us to define a PPP(μ) even if μ is only σ-finite. Indeed, σ-finiteness means that $\mathscr{X}$ admits a partition into a countable family of disjoint measurable sets of finite μ-measure. We now define N by (7.18) on each part of the partition using its own K and X_i's. The resulting random measure is then σ-finite using the same partition as μ.

In view of our motivation to introduce the Poisson point process as a model of noise, a slight problem with the above is that N does not have zero mean. This is fixed by:

Definition 7.4 (Compensated Poisson process) *Let $\{N(A)\colon A\in\Sigma\}$ be a Poisson point process on $\mathscr{X}$ with finite intensity measure μ. Then $\{\widehat{N}(A)\colon A\in\Sigma\}$ for*

$$\widehat{N}(A) := N(A) - \mu(A) \tag{7.9}$$

is the associated compensated Poisson process.

An example of the above setting is the homogeneous Poisson process $\{N_t\colon t\in[0,\infty)\}$ from (3.24) where, abusing notation slightly, $N_t = N([0,t])$ for $\{N(A)\colon A\in\mathbb{B}([0,\infty))\}$ a Poisson point process on $[0,\infty)$ with intensity given by the Lebesgue measure. The compensated point process then evaluates on $[0,t]$ to

$$\widehat{N}_t := N_t - t. \tag{7.10}$$

Thanks to (7.4), the increments of $\widehat{N}$ over non-overlapping interval are independent with zero mean, as desired.

The Poisson point process can be used to define other random objects. For instances, given a PPP(μ) N and a function $f\colon\mathscr{X}\to\mathbb{R}$, the integral

$$Z := \int f(x)N(\mathrm{d}x) \tag{7.11}$$

defines a so called *compound Poisson random variable*. (This uses that N is actually a random measure. One needs to check that Z a measurable map of the underlying probability space to $\mathbb{R}$.) Typically, $\mathscr{X}=\mathbb{R}$ and $f(x)=x$ so we are "summing" the positions of the sample points of N.

One can generalize this trick to define a stochastic process. Indeed, given a Poisson point process on $\mathbb{R}\times[0,\infty)$ with intensity given by a product of a finite measure μ and the Lebesgue measure, we define

$$Z_t := \int_{\mathbb{R}\times[0,t]} x\,N(\mathrm{d}x\mathrm{d}t) \tag{7.12}$$

which (as discussed in 275B) is well defined and finite a.s. whenever $\int|x|\mu(\mathrm{d}x)<\infty$. The trajectory of $t\mapsto Z_t$ is then composed of jumps that "arrive" at random times that are independent Exponentials with parameter $\mu(\mathbb{R})$ with the jumps distributed according to the normalized measure μ. Note that $t\mapsto Z_t$ has right-continuous paths.

7.2 White noise.

The fact that $t\mapsto Z_t$ is not continuous may be undesirable in applications. As it turns out, examples arising from the following constructions are somewhat better behaved (see, for instance, Lemma 7.8):

Definition 7.5 (White noise) *Let $(\mathscr{X},\Sigma,\mu)$ be a measure space with $\mu(\mathscr{X})<\infty$. A (Gaussian) white noise on $(\mathscr{X},\Sigma,\mu)$ is a multivariate Gaussian process $\{W(A)\colon A\in\Sigma\}$ such that*

$$\forall\{A_i\}_{i\in\mathbb{N}}\in\Sigma\colon\ \textit{disjoint}\ \Rightarrow\ \begin{cases}\{W(A_i)\}_{i=1}\ \textit{independent}\\ W\Big(\bigcup_{i\in\mathbb{N}}A_i\Big)=\sum_{i\in\mathbb{N}}W(A_i)\quad\textit{a.s.}\end{cases} \tag{7.13}$$

and

$$\forall A \in \Sigma\colon\ \ W(A) = \mathcal{N}\big(0, \mu(A)\big) \tag{7.14}$$

We start by noting that textbook definitions often specify the above conditions by one simple statement:

$$\forall A, B \in \Sigma\colon \quad E\big[W(A)\big] = 0 \ \wedge\ E\big[W(A)W(B)\big] = \mu(A \cap B) \tag{7.15}$$

That such a Gaussian process can even be considered follows from the observation that

$$\sum_{i,j=1}^{n} a_i a_j \mu(A_i \cap A_j) = \int \bigg|\sum_{i=1}^{n} a_k 1_{A_i}\bigg|^2 \mathrm{d}\mu \geqslant 0 \tag{7.16}$$

which implies that $A, B \mapsto \mu(A \cap B)$ is symmetric and positive semi-definite. Standard arguments from linear algebra then show that $\{\mu(A_i \cap A_j)\}_{i,j=1}^{n}$ is a covariance matrix of an n-dimensional multivariate Gaussian.

We now easily check that a white noise according to Definition 7.5 is a Gaussian process satisfying (7.15). To see that a description based on (7.15) is actually equivalent to the above is now seen from:

Lemma 7.6 *Suppose $\{W(A)\colon A \in \Sigma\}$ is a centered Gaussian process indexed by elements of the σ-algebra Σ such that* (7.15) *holds. Then*

$$W(\varnothing) = 0 \quad \text{a.s.} \tag{7.17}$$

and for all $\{A_n\}_{n\geqslant 0} \subseteq \Sigma$ (pairwise) disjoint,

$$W\Big(\bigcup_{n\geqslant 1} A_n\Big) = \sum_{n\geqslant 1} W(A_n) \quad \text{a.s.} \tag{7.18}$$

where the sum converges in L^2 and a.s.

Proof. Suppose that $\{W(A)\colon A \in \Sigma\}$ is centered Gaussian and (7.15) holds. Then

$$E\big[W(\varnothing)^2\big] = \mathrm{Var}\big(W(\varnothing)\big) + \big[EW(\varnothing)\big]^2 = \mu(\varnothing) + 0 = 0 \tag{7.19}$$

we get (7.17). For (7.18), pick $\{A_n\}_{n\geqslant 1}$ disjoint and abbreviate

$$A := \bigcup_{n\geqslant 1} A_n \tag{7.20}$$

Then (7.15) gives that $\{W(A_n)\}_{n\geqslant 1}$ are independent, zero-mean with

$$\sum_{n\geqslant 1} E\big[W(A_n)^2\big] = \sum_{n\geqslant 1} \mu(A_n) = \mu(A) \leqslant \mu(\mathscr{X}) < \infty. \tag{7.21}$$

The Paley-Zygmund Theorem on convergence of random infinite series (which drives the proof, and is a special case, of the Kolmogorov 3-series Theorem) then implies that the sum on the right of (7.18) converges a.s. and in L^2. Furthermore, the L^2-convergence permits us to write

$$
\begin{aligned}
E\Bigl[\Bigl(W(A)-&\sum_{n\geqslant 1} W(A_n)\Bigr)^2\Bigr] \\
&= E\bigl[W(A)^2\bigr] - 2\sum_{n\geqslant 1} E\bigl[W(A)W(A_n)\bigr] + \sum_{m,n\geqslant 1} E\bigl[W(A_n)W(A_m)\bigr] \\
&= \mu(A) - 2\sum_{n\geqslant 1}\mu(A\cap A_n) + \sum_{m,n\geqslant 1}\mu(A_n\cap A_m) \\
&= \mu(A) - 2\sum_{n\geqslant 1}\mu(A_n) + \sum_{n\geqslant 1}\mu(A_n) = \mu(A) - \sum_{n\geqslant 1}\mu(A_n) = 0
\end{aligned}
\tag{7.22}
$$

thus proving a.s. equality in (7.17) □

Our next task will be the existence of such a process:

Lemma 7.7 *A white noise process exists for each finite measure space $(\mathscr{X},\Sigma,\mu)$.*

Proof. As note above, $\{\mu(A\cap B)\}_{A,B\in\Sigma}$ is a covariance kernel. This means that one can define the finite dimensional distributions of $\{W(A_i)\colon i=1,\dots,n\}$, for any choice of $A_1,\dots,A_n\in\Sigma$, as the law of a multivariate centered normal with covariance matrix $\{\mu(A_i\cap A_j)\}_{i,j=1}^n$. Since the same covariance is used, these finite-dimensional distributions are automatically consistent and so (being defined on $\mathbb{R}$, which is standard Borel under Euclidean metric) the Kolmogorov Extension Theorem gives the existence of a probability space and random variables $\{W(A)\colon A\in\Sigma\}$ with the desired Gaussian law.

Clearly, if $A_1,\dots,A_n\in\Sigma$ are disjoint, then

$$
\forall i\neq j\colon\ \operatorname{Cov}(W(A_i),W(A_j)) = \mu(A_i\cap A_j) = 0 \tag{7.23}
$$

The fact that for multivariate Gaussians uncorrelated implies independent then gives that $W(A_1),\dots,W(A_n)$ are independent. Lemma 7.6 then checks the second condition in (7.13). The condition (7.14) was built into the definition of W. □

As for the Poisson point process, the requirement of finiteness of μ is actually not strictly necessary. Indeed, even if μ is infinite, the above definitely constructs $W(A)$ for each $A\in\Sigma$ with $\mu(A)<\infty$ with the "additivity mod null set" restricted to sets whose total mass is finite. With that proviso in mind, we note the following example:

Lemma 7.8 *Let W be a white noise process on measure space $([0,\infty),\mathcal{B}([0,\infty)),\lambda)$, where λ is the Lebesgue measure. Then $\{W([0,t])\colon t\in[0,\infty)\}$ has the finite-dimensional distributions of standard Brownian motion.*

Proof. It is clear that $t\mapsto W([0,t])$ is a Gaussian process with mean zero and

$$
\operatorname{Cov}\bigl(W([0,t]),W([0,s])\bigr) = \lambda\bigl([0,t]\cap[0,s]\bigr) = t\wedge s. \tag{7.24}
$$

The claim now follows from Lemma 5.2. □

7.3 Paley-Wiener integral.

The identity (7.18) may appear to suggest that $A \mapsto W(A)$ is a signed measure. Unfortunately, unlike the Poisson point process which is a measure proper, this is not the case for the white noise since the implicit null set in (7.18) depends on $\{A_n\}_{n\in\mathbb{N}}$. Still, the absence of σ-additivity does not prevent us from defining "integrals" with respect to this "measure." As in the ordinary Lebesgue integration theory, a starting point is to define the integral of simple functions:

Lemma 7.9 *Let $\{W(A)\colon A \in \Sigma\}$ be a white noise on a measure space $(\mathscr{X}, \Sigma, \mu)$. For any $f\colon \mathscr{X} \to \mathbb{R}$ simple with representation $f := \sum_{i=1}^n a_i 1_{A_i}$ set*

$$\int f \mathrm{d}W := \sum_{i=1}^n a_i W(A_i) \tag{7.25}$$

Then the result does not depend on a representation (modulo modifications on null sets) and, moreover, we have

$$E\left(\left(\int f \mathrm{d}W\right)^2\right) = \int f^2 \mathrm{d}\mu. \tag{7.26}$$

Proof. The independence of the representation is a consequence of a.s. finite additivity of $A \mapsto W(A)$. The equality (7.26) is then proved via (7.15). □

The key point of (7.26) is that it can be viewed as the statement that the map

$$f \mapsto \int f \mathrm{d}W, \tag{7.27}$$

defined presently just for simple functions, is an isometry from $L^2(\mu)$ to L^2 associated with the underlying probability space. As a consequence we get:

Lemma 7.10 *The map* (7.27) *extends continuously to all $f \in L^2$. The identity* (7.26) *continues to hold for the extension.*

The resulting "integral" $\int f \mathrm{d}W$ is called the *Paley-Wiener integral*. We emphasize that this is generally not an integral in the ordinary sense, because W is not a signed measure. But many of the properties of the integral remain in place, although proofs have to be based on L^2-calculus, rather than a.s. limit techniques.