

1 Introduction

The goal of this post is to introduce the theory of *stationary random fields*. In spatio-temporal statistics the typical setup is one wants to model some random quantity X which takes different values at different positions in time and space. That is, one has a collection of random variables $\{X(u)\}_{u \in I}$ over some index I . Typically I is some subset of the temporal domain \mathbb{R}_+ (temporal fields), of a spatial domain \mathbb{R}^d (spatial fields) or a subset of their product $\mathbb{R}_+ \times \mathbb{R}^d$ (space-time fields). But it can also be something more exotic like a manifold or a Hilbert space.

As for what $X(u)$ actually represents it could be anything that interests us and that fundamentally has some uncertainty. For example X could be a weather field, the concentration of a chemical in the stream next to your house, the amount of moose in the Amazon rainforest or even be elements of a Hilbert space.

In this series of posts we'll show the relationship between stationary fields and Fourier analysis through Bochner's theorem and the spectral theorem (links). Furthermore we'll show how differential equations (both SDEs and SPDEs) with constant coefficients have as solutions stationary fields. This fact is very useful both for modeling and numerical purposes. Ok enough chit-chat, time for some definitions that we'll use throughout these posts.

1.1 Preliminary definitions and notation

We consider in what follows a complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$. We will also fix a Banach space E and a complex Hilbert space H with inner product $\langle \cdot, \cdot \rangle$.

Definition 1. *Given a measurable space (N, \mathcal{N}) , we say that $X : \Omega \rightarrow N$ is a N -valued random variable if X is \mathcal{F}/\mathcal{N} measurable.*

The above definition depends on the sigma-algebra \mathcal{N} placed on N . In our work, N will only ever be a topological space. In this case, we will implicitly assume that $\mathcal{N} = \mathcal{B}(N)$ is the Borel sigma-algebra (link) on N .

Definition 2. *Given a measurable space (N, \mathcal{N}) and an arbitrary set I we*

say that a N valued random field with index set I is a collection $\{X(u)\}_{u \in I}$ of N -valued random variables.

We will abbreviate the above to say that $X : I \rightarrow N$ is a (random) field. Furthermore, in the case where the I and N are understood to be fixed, we will simply say that X is a field. We also note that the above measurability condition is equivalent to requiring that X is a N^I valued random variable (see for example [1] page 45).

Given a measure space (M, \mathcal{M}, μ) and a Banach space E , we write $L^p(M \rightarrow E, \mu)$ for the space of equivalence classes of p -integrable functions to E with the norm

$$\|f\|_{L^p(M \rightarrow E)} := \left(\int_{\Omega} \|f\|_E^p d\mu \right)^{\frac{1}{p}}.$$

In practice, we will almost always take $E = \mathbb{R}^n$ or $E = \mathbb{C}^n$. In the case where $E = \mathbb{R}$ we will abbreviate $L^p(M \rightarrow \mathbb{R}, \mu)$ to $L^p(M, \mu)$. Furthermore, if the measure space is $(\Omega, \mathcal{F}, \mathbb{P})$ or $(\Omega, \mathcal{B}(\mathbb{R}^d), \lambda^d)$ where λ^d is the Lebesgue measure on \mathbb{R}^d we will not specify the measure. Writing for example

$$L^p(\mathbb{R}^d \rightarrow \mathbb{C}^n) := L^p(\mathbb{R}^d \rightarrow \mathbb{C}^n, \lambda^d); \quad L^p(\Omega \rightarrow \mathbb{C}^n) := L^p(\Omega \rightarrow \mathbb{C}^n, \mathbb{P}).$$

Definition 3. Given a random field $X : I \rightarrow E$, we say that X is a p -integrable field if $X(u) \in L^p(\Omega \rightarrow E)$ for all $u \in I$.

Finally, given $f \in L^1(\mathbb{R}^d \rightarrow \mathbb{C}^n)$, we will write

$$\hat{f}(\omega) := \left(\frac{1}{2\pi} \right)^d \int_{\mathbb{R}^d} f(u) e^{i\omega \cdot u} du; \quad \check{f}(\omega) := \int_{\mathbb{R}^d} f(u) e^{i\omega \cdot u} du, \quad (1)$$

We recall that the Fourier transform and its inverse extend continuously to isometries on $L^2(\mathbb{R}^d \rightarrow \mathbb{C}^n)$ (which we denote with the same notation $\hat{\cdot}, \check{\cdot}$) with

$$\check{\check{f}} = \hat{\hat{f}} = f \quad \forall f \in L^2(\mathbb{R}^d \rightarrow \mathbb{C}^n). \quad (2)$$

See for example [2] pages 222-226 for a proof.

2 The 1 dimensional case

To begin with, we will consider a \mathbb{C} valued random field X with index set \mathbb{R}^d and then build the multidimensional case of \mathbb{C}^n and even Hilbert valued random field from there.

Definition 4. *Given a square-integrable random field $X : \mathbb{R}^d \rightarrow \mathbb{C}$ we define the covariance function of X as*

$$C(u, v) := \mathbb{E}[X(u)\overline{X(v)}] - \mathbb{E}[X(u)]\overline{\mathbb{E}[X(v)]}$$

Definition 5. *Given a square integrable random field $X : \mathbb{R}^d \rightarrow \mathbb{C}$ we say that X is weakly stationary if for all $u, v \in \mathbb{R}^d$*

$$\mathbb{E}[X(u)] = \mathbb{E}[X(v)]; \quad r(u - v) = C(u, v).$$

For some function $r : \mathbb{R} \rightarrow \mathbb{C}$. By abuse of notation, we will also call r the covariance function of X .

That is, a field is weakly stationary if the covariance between the realizations of the field at two different points only depends on the difference between the points.

The use of weak stationarity is so widespread that it is usually abbreviated to just saying X is stationary and we will use that convention here. That said, a calculation shows that for any (weakly) stationary stochastic field its covariance function must be a Hermitian function, that is,

$$\overline{r(u)} = r(-u).$$

As we will see, the covariance function is one of the fundamental objects of stationary fields, we now discuss its properties.

Definition 6. *A linear operator $T : H \rightarrow H$ of a Hilbert space is said to be positive semi-definite if*

$$\langle Tx, x \rangle \geq 0 \quad \forall x \in H.$$

Where x^\dagger is the adjoint transpose of x .

Since every operator between finite dimensional vector spaces is given by a matrix we use the following definition

Definition 7. We say that a matrix $A \in \mathbb{C}^{n \times n}$ is positive and semi-definite if

$$\langle Ax, x \rangle = x^\dagger Ax \geq 0 \quad \forall x \in \mathbb{C}^n.$$

One example of a positive semi-definite function is the covariance matrix of a \mathbb{C}^n valued random variable Y . That is,

$$A_{ij} = \text{Cov}(Y_i, \overline{Y_j}) = (YY^\dagger - \mathbb{E}[Y]\mathbb{E}[Y^\dagger])_{ij}$$

is positive definite as it holds that given $x \in \mathbb{C}^n$

$$x^\dagger (\mathbb{E}[YY^\dagger] - \mathbb{E}[Y]\mathbb{E}[Y^\dagger])x = \mathbb{E}[\|Y^\dagger x\|^2] - \mathbb{E}[(Y^\dagger x)^\dagger] \mathbb{E}[Y^\dagger x] \geq 0.$$

Where the last inequality is due to Cauchy-Schwartz. In particular, the covariance matrix $(C(u_i, u_j))_{i,j=1}^n$ of any stochastic field is positive semi-definite for all $n \in \mathbb{N}$. If the field is stationary then, by definition,

$$r(u_i - u_j) = C(u_i, u_j),$$

so $(r(u_i - u_j))_{i,j=1}^n$ is positive semi-definite for any $n \in \mathbb{N}$ and for any $u_1, \dots, u_n \in \mathbb{R}^d$. This leads to the following definition.

Definition 8. We say that $r : \mathbb{R}^d \rightarrow \mathbb{C}$ is positive semi-definite if the matrix $(r(u_i - u_j))_{i,j=1}^n$ is positive semi-definite for any $n \in \mathbb{N}$ and for any $u_1, \dots, u_n \in \mathbb{R}^d$. That is, if

$$\sum_{j,k=1}^n r(u_j - u_k) a_j \bar{a}_k \geq 0 \quad \forall a_j \in \mathbb{C}, u_j \in \mathbb{R}^d.$$

Since the above sum must be real, a small reasoning shows that any positive semi-definite function must be Hermitian. Thus, it is reasonable to inquire whether every positive semi-definite function is the covariance function of some field.

Proposition 1. There is a 1 to 1 equivalence between positive semi-definite functions and covariance functions of weakly stationary fields.

Proof. We have already seen that every covariance function is positive semi-definite, so it remains to check that a positive semi-definite function $r(u)$ corresponds to the covariance function of some weakly stationary field. We

can choose the process to be Gaussian. This follows by the fact that given any positive semi-definite Hermitian matrix A , we can always find a Gaussian variable with covariance matrix A . This allows us to construct the desired family of finite dimensional distributions $\{X_j\}_{j \in J}$, and then an application of Kolmogorov's extension theorem (link) concludes the proof. See [3] page 72 for the details. \square

There, is a nice characterization of the positive semi-definite functions first proved by Bochner in [4]

Theorem 1 (Bochner's theorem). *A function $r : \mathbb{R}^d \rightarrow \mathbb{C}$ is a continuous positive semi-definite function if and only if for some finite measure $\mu : \mathcal{B}(\mathbb{R}^d) \rightarrow \mathbb{R}_+$*

$$r(u) = \int_{\mathbb{R}^d} e^{i\omega \cdot u} d\mu(\omega).$$

Furthermore, μ is unique.

We note that the implication part of the theorem can be seen from the fact

$$\sum_{j,k=1}^n a_j \bar{a}_k \int_{\mathbb{R}^d} e^{i\omega \cdot u_j} e^{-i\omega \cdot u_k} d\mu(\omega) = \int_{\mathbb{R}^d} \left| \sum_{j=1}^n a_j e^{i\omega \cdot u_j} \right|^2 d\mu(\omega) \geq 0.$$

The converse is however more technical and follows from a more general result on harmonic analysis of groups. We will discuss this in more detail later on.

In the language of probability theory, another way of stating the above relationship is to say that r is the *characteristic function* of μ and to write $r = \hat{\mu}$. We note that μ is a *Lebesgue-Stieltjes measure* (that is, a measure on $\mathcal{B}(\mathbb{R}^d)$ which is finite on bounded intervals) and as a result the function $F : \mathbb{R}^d \rightarrow \mathbb{R}$ defined by

$$F(\omega) := \mu(-\infty, \omega].$$

Is, save additive constants, the unique *distribution function* on \mathbb{R}^d (that is, F is right continuous, non-decreasing and bounded) such that $\mu = dF$. See for example [5] page 30 for a proof of this fact. We call respectively the measure μ and distribution function F appearing in Bochner's theorem the *spectral measure* and *spectral distribution* of X . In the case where r is integrable, μ will have a density.

Proposition 2 (Fourier inversion). *Let $r \in L^1(\mathbb{R}^d \rightarrow \mathbb{C})$ be a continuous positive definite function with spectral distribution dF . Then $S(\omega) := \partial_1 \cdots \partial_d F(\omega)$ exists almost everywhere, and we have*

$$S(\omega) = \left(\frac{1}{2\pi} \right)^d \int_{\mathbb{R}^d} r(u) e^{-i\omega \cdot u} du; \quad r(u) = \int_{\mathbb{R}^d} S(\omega) e^{i\omega \cdot u} d\omega.$$

The proof is based on a classical result of probability called the inversion formula (see [5] page 292).

In the case of Proposition 2 above it is customary to say that S is the *spectral density* of X . We observe that the condition for the existence of the spectral density is quite weak. Essentially, all that is required is that the relationship between the field at points that are far decays to 0 at a fast enough rate. For completeness, we mention the following equivalent definition which may be found in some references

Lemma 1. *A continuous function $r : \mathbb{R}^d \rightarrow \mathbb{C}$ is positive semi-definite if and only if for all $f \in L^1(\mathbb{R}^d \rightarrow \mathbb{C})$*

$$\int_{\mathbb{R}^d} \int_{\mathbb{R}^d} r(x-y) f(x) \overline{f(y)} dx dy \geq 0.$$

Proof. Let us denote the above integral by $I(f)$. First, we show that $I(f)$ is well-defined and continuous as a function of $f \in L^1(\mathbb{R}^d \rightarrow \mathbb{C})$. By Bochner's theorem (Theorem 1) r takes a maximum at 0 so

$$|I(f)| \leq r(0) \left| \int_{\mathbb{R}^d} f(x) dx \right| \left| \int_{\mathbb{R}^d} \overline{f(y)} dy \right| \leq r(0) \|f\|_{L^1(\mathbb{R}^d \rightarrow \mathbb{C})}^2$$

To prove the implication, one can first work with $f \in C_c(\mathbb{R}^d \rightarrow \mathbb{C})$. In this case, the above integral can be approximated by a sum over increasingly fine partitions

$$I(f) = \lim_{N \rightarrow \infty} \sum_{j,k=1}^{\infty} \frac{1}{(x_{k+1}^{(N)} - x_k^{(N)})(x_{j+1}^{(N)} - x_j^{(N)})} r(x_k^{(N)} - x_j^{(N)}) f(x_k^{(N)}) \overline{f(x_j^{(N)})}.$$

Since all these sums are positive, so must be the integral. Then we can use the density of $C_c(\mathbb{R}^d \rightarrow \mathbb{C})$ in $L^1(\mathbb{R}^d \rightarrow \mathbb{C})$ to get the general result. For the

converse, given $a_1, \dots, a_n \in \mathbb{C}$, one can take $f_N \in \mathbb{C}_c(\mathbb{R}^d \rightarrow \mathbb{C})$ converging to $\sum_j a_j \delta(x_j)$ in the sense that

$$\lim_{N \rightarrow \infty} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} r(x-y) f_N(x) \overline{f_N(y)} dx dy = \sum_{j,k=1}^n r(u_j - u_k) a_j \bar{a}_k.$$

Since, by hypothesis, the above integrals are greater than 0 is their limit. This concludes the proof. \square

The regularity of r is intimately linked to the (mean square) continuity of X .

Lemma 2. *Let $X : \mathbb{R}^d \rightarrow \mathbb{C}$ be a stationary field with covariance function r . Then if r is continuous X is mean square continuous. That is*

$$\lim_{h \rightarrow 0} \mathbb{E}[|X(u+h) - X(u)|^2] = 0, \quad \forall u \in \mathbb{R}^d.$$

Proof. We have that, by a calculation,

$$\begin{aligned} \mathbb{E}[|X(u+h) - X(u)|^2] &= \mathbb{E}[|X(h)|^2] + \mathbb{E}[|X(h)|^2] - \mathbb{E}[X(u+h)\overline{X(u)}] \\ &\quad - \mathbb{E}[X(u)\overline{X(u+h)}] = 2(r(0) - \operatorname{Re}(r(h))). \end{aligned}$$

If r is continuous, since by definition $r(0)$ is positive we have that the above goes to 0 as h goes to 0, which concludes the proof. \square

The result has a converse (see [3] page 34), but we shall not need it. Our future results will all be based on Bochner's theorem and thus require the continuity of X . In what remains of the work, we assume that all stationary fields have continuous covariance, or equivalently are mean square continuous.

We next construct the spectral process dZ of X . This, along with the spectral distribution, dF form the most crucial objects in the spectral theory of stationary fields.

Theorem 2 (The spectral theorem). *If $X : \mathbb{R}^d \rightarrow \mathbb{C}$ is a stationary field with mean zero and spectral distribution $F(\omega)$, there exists a unique generalized field dZ over $L^2(\mathbb{R}^d \rightarrow \mathbb{C}, dF)$ such that*

$$X(u) = dZ(e^{iu \cdot}), \quad \forall u \in \mathbb{R}^d \tag{3}$$

Furthermore, dZ defines a bijective isometry

$$dZ : L^2(\mathbb{R}^d \rightarrow \mathbb{C}, dF) \rightarrow \overline{\text{span}(\{X(u)\}_{u \in \mathbb{R}^d})} \subset L^2(\Omega \rightarrow \mathbb{C}).$$

In particular, for any $f, g \in L^2(\mathbb{R}^d \rightarrow \mathbb{C}, dF)$,

$$\mathbb{E}[dZ(f)] = 0; \quad \mathbb{E}[dZ(f)\overline{dZ(g)}] = \int_{\mathbb{R}^d} f\bar{g}dF. \quad (4)$$

Proof. We will first prove the existence of Z and then use it to build dZ . Let us define given $u \in \mathbb{R}^d$

$$\varphi(X(u)) := e^{iu \cdot}. \quad (5)$$

Where $e^{iu \cdot}$ is a function of \cdot and is in $L^2(\mathbb{R}^d \rightarrow \mathbb{C}, dF)$ as dF is a bounded measure. We then extend this definition linearly to the span of $\{X(u)\}_{u \in \mathbb{R}^d}$ by setting

$$\varphi\left(\sum_{j=1}^N \lambda_j X(u_j)\right) := \sum_{j=1}^N \lambda_j e^{iu_j \cdot} \quad \forall N \in \mathbb{N}, u_j \in \mathbb{R}^d.$$

Our next step is to show that φ defines an isometry

$$\varphi : \text{span}(\{X(u)\}_{u \in \mathbb{R}^d}) \subset L^2(\Omega \rightarrow \mathbb{C}) \rightarrow L^2(\mathbb{R}^d \rightarrow \mathbb{C}, dF).$$

Let $Y := \sum_{j=1}^N \lambda_j X(u_j)$, then

$$\begin{aligned} \|Y\|_{L^2(\Omega \rightarrow \mathbb{C})}^2 &= \left\| \sum_{j=1}^N \lambda_j X(u_j) \right\|_{L^2(\Omega \rightarrow \mathbb{C})}^2 = \sum_{j,k=1}^N \lambda_j \bar{\lambda}_k \mathbb{E}[X(u_j) \overline{X(u_k)}] \\ &= \sum_{j,k=1}^N \lambda_j \bar{\lambda}_k r(u_j - u_k) = \sum_{j,k=1}^N \lambda_j \bar{\lambda}_k \int_{\mathbb{R}^d} e^{i\omega \cdot (u_j - u_k)} dF(w) \\ &= \int_{\mathbb{R}^d} \left| \sum_{j=1}^N \lambda_j e^{i\omega \cdot u_j} \right|^2 dF(w) = \|\varphi(Y)\|_{L^2(\mathbb{R}^d \rightarrow \mathbb{C}, dF)}^2 \end{aligned}$$

By the completeness of $L^2(\mathbb{R}^d \rightarrow \mathbb{C}, dF)$ and since φ is an isometry, we can extend φ to a mapping

$$\varphi : \overline{\text{Span}(\{X(u)\}_{u \in \mathbb{R}^d})} \rightarrow L^2(\mathbb{R}^d \rightarrow \mathbb{C}, dF).$$

Furthermore, by the Stone-Weierstrass theorem, the span on $\{e^{iu}\}_{u \in \mathbb{R}^d}$ (which by construction is in the image of φ) is dense in $L^2(\mathbb{R}^d \rightarrow \mathbb{C}, dF)$. In consequence, φ is surjective (and thus bijective) and we can now define

$$dZ(f) := \varphi^{-1}(f) \quad \forall f \in L^2(\mathbb{R}^d \rightarrow \mathbb{C}, dF).$$

Property (3) follows from the construction of φ (see equation (5)). The first part of (4) follows from the fact that φ is valued in the closure of the span of $X(u)$ (which is of mean zero). The second part of (4) is because dZ is an isometry, and thus also conserves the inner product. Finally, the uniqueness of dZ follows from the fact that, by property (3), dZ is determined on a dense subset of $L^2(\mathbb{R}^d \rightarrow \mathbb{C}, dF)$ and the continuity of property (4). \square

In the literature ([6] page 98, [3] page 89) one may often find the notation

$$\int f(\omega) dZ(\omega) := dZ(f), \quad f \in L^2(\mathbb{R}^d \rightarrow \mathbb{C}, dF).$$

We will also use this notation, which corresponds with identifying dZ with some kind of stochastic integration. To motivate this idea, let us set

$$Z(\omega) := dZ(1_{(-\infty, \omega]}), \tag{6}$$

and consider for simplicity the 1 dimensional case $d = 1$. Then, given any simple function

$$f(\omega) = \sum_{j=1}^n \lambda_j 1_{(\omega_j, \omega_{j+1}]}(\omega),$$

we have that by the linearity of dZ

$$dZ(f) = \sum_{j=1}^n \lambda_j dZ(1_{(\omega_j, \omega_{j+1}]}) = \sum_{j=1}^n \lambda_j (Z(\omega_{j+1}) - Z(\omega_j)).$$

Where the above sum is identical in form to that of Riemann-Lebesgue integration of f against dZ . Furthermore, by the density of simple function in $L^2(\mathbb{R}, dF)$, given any $f \in L^2(\mathbb{R}, dF)$ we may take a sequence of simple functions

$$f_n(\omega) = \sum_{j=1}^n \lambda_j 1_{(\omega_j^{(n)}, \omega_{j+1}^{(n)}]}(\omega),$$

converging to f . Then, since dZ is an isometry (and thus continuous) we have that

$$dZ(f) = \lim_{n \rightarrow \infty} dZ(f_n) = \lim_{n \rightarrow \infty} \sum_{j=1}^n f(\omega_j^{(n)}) \left(Z(\omega_{j+1}^{(n)}) - Z(\omega_j^{(n)}) \right) \quad (7)$$

So once more we recover $dZ(f)$ as a limit of stochastic integrals. This kind of construction mirrors the construction of the Itô integral. Where in our case dZ takes the place of white noise dW . This also justifies why some authors call dZ a *dF noise on \mathbb{R}^d* ([7] page 105).

Observation 1. *The spectral theorem means that in the future we will usually work with fields X with mean 0. However, many of the results can be generalized without difficulty to the case where $\mathbb{E}[X] \neq 0$ by noting that*

$$X(u) = \mathbb{E}[X] + \int_{\mathbb{R}^d} e^{i\omega \cdot u} dZ(\omega).$$

Where dZ is now the spectral process of $X - \mathbb{E}[X]$. Equivalently, one can also write

$$X = \int_{\mathbb{R}^d} e^{i\omega \cdot u} \widetilde{dZ}(\omega).$$

Where now $\widetilde{dZ} = dZ + \mathbb{E}[X]\delta_0$ will no longer be a *dF noise* but rather verify for any given $f \in L^2(\mathbb{R}^d \rightarrow \mathbb{C})$.

$$\mathbb{E}[\widetilde{dZ}(f)] = \mathbb{E}[X]f(0); \quad \mathbb{E}[(\widetilde{dZ}(f))^2] = \|f\|_{L^2(\mathbb{R}^d \rightarrow \mathbb{C})}^2 + \mathbb{E}[X]^2 f(0)$$

We now prove an interesting corollary of the spectral representation theorem.

Corollary 1. *Let X be a mean zero, a stationary field with spectral distribution and spectral process dF, dZ respectively. Then $\{X(u)\}_{u \in \mathbb{R}^d}$ is Gaussian if and only if.*

$$dZ(f) \sim \mathcal{N}\left(0, \|f\|_{L^2(\mathbb{R}^d \rightarrow \mathbb{C}, dF)}^2\right) \quad \forall f \in L^2(\mathbb{R}^d \rightarrow \mathbb{C}, dF).$$

Proof. We recall that, by the basic theory of Gaussian random variables, if Y_n is a sequence of Gaussian random variables converging in distribution to

Y then Y is also Gaussian. Furthermore, as we showed in Theorem 2 (the spectral theorem) for each f ,

$$dZ(f) \in \overline{\text{Span}(\{X(u)\}_{u \in \mathbb{R}^d})} \subset L^2(\mathbb{R}^d \rightarrow C, dF).$$

Since convergence in L^p implies convergence in distribution we deduce that if X is Gaussian then $dZ(f)$ is also Gaussian. Reciprocally, if dZ is Gaussian, then necessarily $X(u) = dZ(e^{i\omega \cdot u})$ is Gaussian. Finally, the calculation of the variance of dZ follows once more using that dZ is an isometry. \square

In the conditions of the above corollary, we will say that dZ is a *Gaussian dF noise on \mathbb{R}^d* . Additionally, we also note that $Z(\omega)$ defined as in (6) must be Gaussian as follows by setting $f = 1_{(-\infty, \omega]}$. Next, we prove a stochastic version of Fubini's theorem, which we will use later in Subsection ??.

Proposition 3 (Stochastic Fubini). *Given $f \in L^1(\mathbb{R}^d \rightarrow L^2(\mathbb{R}^d \rightarrow \mathbb{C}, dF))$ and a stochastic dF noise Z , it holds that*

$$\int_{\mathbb{R}^d} \left(\int_{\mathbb{R}^d} f(v, \omega) dv \right) dZ(\omega) = \int_{\mathbb{R}^d} \left(\int_{\mathbb{R}^d} f(v, \omega) dZ(\omega) \right) dv.$$

Proof. Firstly, a minimum requirement is that both sides be well-defined as functions in $L^2(\Omega \rightarrow \mathbb{C})$. That is, that

$$\int_{\mathbb{R}^d} f(v, \cdot) dv \in L^2(\mathbb{R}^d \rightarrow \mathbb{C}, dF); \quad \int_{\mathbb{R}^d} f(\cdot, \omega) dZ(\omega) \in L^1(\mathbb{R}^d \rightarrow \mathbb{C}).$$

The leftmost inclusion holds by Minkowski's integral inequality as

$$\begin{aligned} \left\| \int_{\mathbb{R}^d} f(v, \cdot) dv \right\|_{L^2(\mathbb{R}^d \rightarrow \mathbb{C}, dF)} &\leq \int_{\mathbb{R}^d} \|f(v, \cdot)\|_{L^2(\mathbb{R}^d \rightarrow \mathbb{C}, dF)} dv \\ &= \|f\|_{L^1(\mathbb{R}^d \rightarrow L^2(\mathbb{R}^d \rightarrow \mathbb{C}, dF))} < \infty \end{aligned}$$

The right-hand side is also well-defined as, by Minkowski's inequality and the isometry of the stochastic integral

$$\begin{aligned} \left\| \left\| \int_{\mathbb{R}^d} f(\cdot, \omega) dZ(\omega) \right\|_{L^1(\mathbb{R}^d \rightarrow \mathbb{C})} \right\|_{L^2(\Omega \rightarrow \mathbb{C})} &\leq \left\| \left\| \int_{\mathbb{R}^d} f(\cdot, \omega) dZ(\omega) \right\|_{L^2(\Omega \rightarrow \mathbb{C})} \right\|_{L^1(\mathbb{R}^d \rightarrow \mathbb{C})} \\ &= \|f\|_{L^1(\mathbb{R}^d \rightarrow L^2(\mathbb{R}^d \rightarrow \mathbb{C}, dF))} < \infty. \quad (8) \end{aligned}$$

To now prove the equality consider the case where $f = 1_{A \times B}$ is the indicator function of some rectangle where $A \in L^1(\mathbb{R}^d \rightarrow \mathbb{C})$ and $B \in L^2(\mathbb{R}^d \rightarrow \mathbb{C}, dF)$. Then we have that

$$\begin{aligned} \int_{\mathbb{R}^d} \left(\int_{\mathbb{R}^d} f(v, \omega) dv \right) dZ(\omega) &= \int_{\mathbb{R}^d} \lambda_d(A) 1_B(\omega) dZ(\omega) = \lambda_d(A) dZ(1_B). \\ \int_{\mathbb{R}^d} \left(\int_{\mathbb{R}^d} f(v, \omega) dZ(\omega) \right) dv &= \int_{\mathbb{R}^d} 1_A(v) dZ(1_B) dv = \lambda_d(A) dZ(1_B). \end{aligned}$$

Where in the above we wrote λ_d for the Lebesgue measure, used the linearity of the Lebesgue and stochastic integrals, and the definition of the stochastic integral. Thus, in this second case, the equality of the proposition holds. Moreover, if f is a linear combination of such indicator functions, an identical reasoning holds. Finally, it is a standard argument to show that f may be approximated by an increasing sequence of linear combinations of rectangles $f_n \leq f$. By applying inequality (8) together with the monotone convergence theorem we get

$$\begin{aligned} \left\| \int_{\mathbb{R}^d} \left(\int_{\mathbb{R}^d} (f - f_n)(v, \omega) dv \right) dZ(\omega) \right\|_{L^2(\Omega \rightarrow \mathbb{C})} &\leq \|f - f_n\|_{L^1(\mathbb{R}^d \rightarrow L^2(\mathbb{R}^d \rightarrow \mathbb{C}, dF))} \rightarrow 0 \\ \left\| \int_{\mathbb{R}^d} \left(\int_{\mathbb{R}^d} (f - f_n)(v, \omega) dZ(\omega) \right) dv \right\|_{L^2(\Omega \rightarrow \mathbb{C})} &\leq \|f - f_n\|_{L^1(\mathbb{R}^d \rightarrow L^2(\mathbb{R}^d \rightarrow \mathbb{C}, dF))} \rightarrow 0. \end{aligned}$$

Since the equality of the proposition holds for f_n by the convergence above we conclude the same for f as desired \square

3 The multidimensional and H valued case

The results of the previous subsection for complex-valued fields extend naturally to the multidimensional case, where X is valued in \mathbb{C}^n . To the best of our knowledge the case where X is valued in some (possibly non-separable) Hilbert space H is not treated in the literature, however, this case follows without too much difficulty from the multidimensional case.

We begin by extending the definition of stationarity.

Definition 9. *Let H be a complex Hilbert space with orthonormal basis $\{e_j\}_{j \in J}$. Then we say that a field X indexed over G and with values in*

H is stationary if it is square-integrable and for all $u, v \in G$ and $j, k \in J$ there are functions $r_{jk} : G \rightarrow H$ verifying

$$\mathbb{E}[X(u)] = \mathbb{E}[X(v)]; \quad r_{jk}(u - v) = \mathbb{E}[X_j(u)\overline{X_k(v)}] - \mathbb{E}[X_j(u)]\overline{\mathbb{E}[X_k(v)]}.$$

Where $X_j := \langle X, e_j \rangle$ is the projection onto element e_j .

We will say that $r : \mathbb{R}^d \rightarrow \mathbb{C}^{J \times J}$ is the *covariance function* of X (with respect to $\{e_j\}_{j \in J}$). First, we show that, as is desirable, stationarity is independent of the basis chosen.

Lemma 3. *Let X be a stationary field valued in a Hilbert space H with orthonormal basis $\{e_j\}_{j \in J}$. Then X is also stationary with respect to any other orthonormal basis $\{v_\beta\}_{\beta \in K}$.*

Proof. We must check that the covariance of the projections onto v_β only depends on the distance. This follows from the fact that, by Plancherel's theorem

$$\begin{aligned} \mathbb{E}[\langle X(u), v_j \rangle \overline{\langle X(v), v_k \rangle}] &= \mathbb{E} \left[\sum_{l \in J} X_l \langle e_l, v_j \rangle \overline{\sum_{m \in J} X_m \langle e_m, v_k \rangle} \right] \\ &= \sum_{l \in J} \sum_{m \in J} \langle e_l, v_j \rangle \overline{\langle e_m, v_k \rangle} r_{lm}(u - v) =: \tilde{r}_{jk}(u - v). \end{aligned}$$

Where the commutation of the sum is justified as, once more by Plancherel, the sums on the right-hand side of the first equality converge in the space $L^2(\Omega \rightarrow \mathbb{C})$. \square

Though stationarity is independent of the basis chosen, the covariance function won't be. In the case where $H = \mathbb{C}^n$, we will always consider the covariance function with respect to the canonical basis on \mathbb{C}^n .

Having set up the theory for general Hilbert spaces, we now somewhat restrict to this case. That is, we consider a stationary field $\{X(u)\}_{u \in \mathbb{R}^d}$ valued in \mathbb{C}^n . The first step in extending the spectral theory of Section 2 is to extend Bochner's theorem. To do so one must be a bit careful as $r_{jk} := \text{Cov}(X_j, X_k)$ is *not* necessarily positive semi-definite for $j \neq k$.

Theorem 3 (Bochner's multidimensional theorem). *If the covariance function of a square-integrable stationary field $\{X(u)\}_{u \in \mathbb{R}^d}$ valued in \mathbb{C}^n is continuous, there exist complex measures $d\mu_{jk}$ such that*

$$r_{jk} = \int_{\mathbb{R}^d} e^{i\omega \cdot u} d\mu_{jk}(\omega), \quad \forall j, k \in 1, \dots, n.$$

Proof. However, given any $a = (a_1, \dots, a_n) \in \mathbb{C}^n$ the random process $Y := a \cdot X$ is a stationary complex-valued field with covariance

$$r_a := \sum_{j,k=1}^n a_j \overline{a_k} r_{jk}.$$

Thus, by the 1-dimensional Bochner's theorem (Theorem 1) we can find some distribution function dF_a such that

$$r_a = \sum_{j,k=1}^n a_j \overline{a_k} r_{jk} = \int_{\mathbb{R}^d} e^{i\omega \cdot u} dF_a(\omega) \quad (9)$$

Given $j, k \in \{1, \dots, n\}$ we define

$$\begin{aligned} a^{(j,k)} &:= (0, \dots, 0, \overset{(j)}{1}, 0, \dots, 0, \overset{(k)}{1}, 0, \dots, 0) \\ b^{(j,k)} &:= (0, \dots, 0, \overset{(j)}{i}, 0, \dots, 0, \overset{(k)}{1}, 0, \dots, 0) \\ c^{(j)} &:= (0, \dots, 0, \overset{(j)}{1}, 0, \dots, 0). \end{aligned}$$

From (9) we obtain that

$$\begin{aligned} r_{jk} + r_{kj} + r_{jj} + r_{kk} &= \int_{\mathbb{R}^d} e^{i\omega \cdot u} dF_{a^{(j,k)}}(\omega); & r_{jj} &= \int_{\mathbb{R}^d} e^{i\omega \cdot u} dF_{c^{(j)}}(\omega) \\ ir_{jk} - ir_{kj} + r_{jj} + r_{kk} &= \int_{\mathbb{R}^d} e^{i\omega \cdot u} dF_{b^{(j,k)}}(\omega); & r_{kk} &= \int_{\mathbb{R}^d} e^{i\omega \cdot u} dF_{c^{(k)}}(\omega). \end{aligned}$$

The above system can be solved to obtain

$$r_{jk} = \int_{\mathbb{R}^d} e^{i\omega \cdot u} dF_{jk}$$

Where dF_{jk} is a complex linear combination of $dF_{a^{(j,k)}}$, $dF_{b^{(j,k)}}$, $dF_{c^{(j)}}$, $dF_{c^{(k)}}$ and is thus a *complex* valued measure (link). \square

From the above, we can quickly obtain Bochner's theorem for fields valued in Hilbert spaces

Theorem 4 (Bochner's theorem for H valued fields). *Let H be a Hilbert space with orthonormal basis $\{e_j\}_{j \in J}$ and let $\{X(u)\}_{u \in \mathbb{R}^d}$ be a stationary H valued field. Then, if the covariance function $r : \mathbb{R}^d \rightarrow \mathbb{C}^{J \times J}$ of X with respect to $\{e_j\}_{j \in J}$ is continuous, there exist complex measures μ_{jk} such that*

$$r_{jk} = \int_{\mathbb{R}^d} e^{i\omega \cdot u} d\mu_{jk}(\omega), \quad \forall j, k \in J.$$

Furthermore, μ_{jj} are bounded real measures for all $j \in \{1, \dots, n\}$.

Proof. Let us fix $j \neq k \in J$ and consider the \mathbb{C}^2 valued field $\{(X_j(u), X_k(u))\}_{u \in \mathbb{R}^d}$. Then by applying Theorem 3 (Bochner's multidimensional theorem) we obtain complex measures μ_{jk} verifying

$$r_{jk} = \int_{\mathbb{R}^d} e^{i\omega \cdot u} d\mu_{jk}(\omega).$$

To see that μ_{jj} are real-valued, it suffices to consider the complex-valued field $\{(X_j(u))\}_{u \in \mathbb{R}^d}$ and once more apply Bochner's theorem. This completes the proof. \square

The analog of the spectral theorem (Theorem 2) is also immediate

Theorem 5 (Spectral theorem for H valued fields). *Let H be a Hilbert space with orthonormal basis $\{e_j\}_{j \in J}$ and let $\{X(u)\}_{u \in \mathbb{R}^d}$ be a stationary H valued field with mean zero and spectral measure μ_{jk} . Then there exists a unique H valued generalized field dZ over $L^2(\mathbb{R}^d \rightarrow \mathbb{C}, dF)$ such that,*

$$X(u) = dZ(e^{iu \cdot}), \quad \forall u \in \mathbb{R}^d \tag{10}$$

Furthermore, dZ defines a bijective isometry

$$dZ : L^2(\mathbb{R}^d \rightarrow \mathbb{C}, \nu) \rightarrow \overline{\text{span}(\{X(u)\}_{u \in \mathbb{R}^d})} \subset L^2(\Omega \rightarrow H).$$

Where $\nu := \sum_{j \in J} \mu_{jj}$. In particular, for any $f, g \in L^2(\mathbb{R}^d \rightarrow \mathbb{C}, \nu)$,

$$\mathbb{E}[dZ(f)] = 0; \quad \mathbb{E}[\langle dZ(f), dZ(g) \rangle] = \int_{\mathbb{R}^d} f \bar{g} \nu. \tag{11}$$

Proof. Since $X_j := \langle X, e_j \rangle$ are stationary we deduce from the spectral theorem for \mathbb{C} valued stationary fields (Theorem 2) that there exist dZ_j verifying the properties of said theorem for X_j . Let us set

$$dZ(f) := \sum_{j \in J} dZ_j(f) e_j \quad \forall f \in L^2(\mathbb{R}^d \rightarrow \mathbb{C}).$$

Then, by the construction of dZ_j ,

$$X(u) = \sum_{j \in J} X_j(u) e_j = \sum_{j \in J} dZ_j(e^{iu \cdot}) e_j = dZ(e^{iu \cdot}).$$

The fact that $dZ(u) \in \overline{\text{span}(\{X(u)\}_{u \in \mathbb{R}^d})} \subset L^2(\Omega \rightarrow H)$ also follows from the construction of dZ and the properties of dZ_j . Finally, the isometry in (11) uses the fact that dZ_j define isometries and Plancherel as,

$$\mathbb{E}[\langle dZ(f), dZ(g) \rangle] = \sum_{j \in J} \mathbb{E}[dZ_j(f) \overline{dZ_j(g)}] = \sum_{j \in J} \int_{\mathbb{R}^d} f \bar{g} d\mu_{jj} = \int_{\mathbb{R}^d} f \bar{g} d\nu.$$

Where ν is well-defined since, as was observed at the end of Theorem 3, μ_{jj} are real measures (as opposed to complex). \square

We conclude this subsection by observing that the spectral theory may also be extended to cases where $\{X(u)\}_{u \in G}$ is a field over some group $(G, +)$. The following discussion is based on [7] pages 119 and [8].

First, we note that the extension of the definition of covariance function and stationarity has no complication. For simplicity, let us now assume that G is a locally compact Abelian group and that X is a stationary field valued in \mathbb{C} . We have that the group of characters

$$\hat{G} := \{ \text{group homomorphisms } \omega : (G, +) \rightarrow (\mathbb{C}, \cdot) : |f(g)| = 1, \quad \forall g \in G \}$$

can also be made into a locally compact Abelian group. Furthermore, one can show the existence of a spectral measure μ and a μ noise dZ such that for all $g \in G$

$$r(g) = \int_{\hat{G}} \omega(g) d\mu(\omega); \quad X(g) = \int_{\hat{G}} \omega(g) dZ(\omega).$$

An example of this theory occurs for example when $G = \mathbb{Z}^d$ (that is, X is a stationary sequence of random variables), then $\hat{G} = \{e^{i\omega \cdot} : \omega \in [-\pi, \pi]^d\}$ and we obtain that for each $k \in \mathbb{Z}^d$

$$r(k) = \int_{[-\pi, \pi]^d} e^{i\omega \cdot k} d\mu(\omega).$$

If on the other hand $G = [-\pi, \pi]^d$ then $\hat{G} = \{e^{ik \cdot} : k \in \mathbb{Z}^d\}$ is discrete and we get that for each $u \in [-\pi, \pi]^d$

$$r(u) = \sum_{k \in \mathbb{Z}^d} e^{ik \cdot u} \mu_k.$$

4 SDES and stationary fields

4.1 Linear filters: Frequency and impulse response

In this subsection, we consider a very useful linear transformation of stationary fields.

Definition 10. *Given a mean zero stationary field $X : \mathbb{R}^d \rightarrow \mathbb{C}$ with spectral measure and process dF, dZ respectively, a linear invariant filter is an operator \mathcal{F}_g of the form*

$$\mathcal{F}_g(X)(u) := \int_{\mathbb{R}^d} g(\omega) e^{i\omega \cdot u} dZ(\omega).$$

where $g \in L^2(\mathbb{R}^d \rightarrow \mathbb{C}, dF)$ is called the frequency response (or transfer function).

By the equality

$$X(u) = \int_{\mathbb{R}^d} e^{i\omega \cdot u} dZ(\omega)$$

it is natural to view dZ as the Fourier transform of X . This is quite similar to the concept of Fourier multipliers appearing in Harmonic analysis. That is, we are modifying our *input field* X by multiplying its Fourier transform by some function g . Thus modifying the frequencies that appear in its representation. Formally, if we write \mathcal{F} for the Fourier transform we have that

$$\mathcal{F}_g(X) = \mathcal{F}^{-1}(g\mathcal{F}(X)) \tag{12}$$

A calculation shows that the following two properties are verified

Lemma 4. *Let \mathcal{F}_g be a linear filter, and consider two stationary fields X, Y such that $X + Y$ is also stationary.*

1. *Linearity:* $\mathcal{F}_g(X + Y) = \mathcal{F}_g(X) + \mathcal{F}_g(Y)$.
2. *Translation invariance:* Given any $v \in \mathbb{R}^d$ we have that $\tau_v \mathcal{F}_g = \mathcal{F}_g \tau_v$ where $\tau_v : \mathbb{R}^d \rightarrow \mathbb{R}^d$ defined by $\tau_v(u) = u + v$ is the translation by v .

Proof. The linearity follows from the fact that, if X, Y have respective spectral processes Z_1, Z_2 , then $X + Y$ has spectral process $Z_1 + Z_2$. The spectral invariance follows from the fact that the spectral representation of the translation $\tau_v X := X(\cdot + v)$ is, just like in Fourier analysis $e^{i\omega v} Z_1$. \square

One can also generalize the definition of a linear filter as any operator that satisfies the two properties above. Though we will not need to do so ourselves. One way of defining a linear translation invariant operator is by setting for $h \in L^1(\mathbb{R}^d \rightarrow \mathbb{C})$

$$Y(u) := \int_{\mathbb{R}^d} h(v) X(u - v) dv. \quad (13)$$

We note that the integral in (13) is well-defined as, by Fubini's theorem,

$$\mathbb{E}[Y] = \int_{\mathbb{R}^d} h(v) \mathbb{E}[X(u - v)] dv = \mu \int_{\mathbb{R}^d} h(v) \leq \mu \|h\|_{L^1(\mathbb{R}^d \rightarrow \mathbb{C})}.$$

Where μ is the constant mean of the field. The terms appearing in (13) respectively: the *response function* Y , the *input function* X , and the *impulse response* h . This last name is motivated by the fact that if $X(u)$ is an impulse $\delta_0(u)$ (yes we know this isn't a stationary field) then $Y(u) = h(u)$.

A crucial fact is that the impulse response h and the frequency response g have an intimate relationship. The frequency response corresponding to h is $(2\pi)^d \hat{h}$, whereas the impulse response corresponding to g is $(2\pi)^{-d} \check{g}$. That is,

$$g = (2\pi)^d \hat{h}; \quad h = \frac{1}{(2\pi)^d} \check{g}.$$

We encode this in the result below.

Proposition 4. *Given a stationary field $X : \mathbb{R}^d \rightarrow \mathbb{C}$ with mean zero and $h \in L^1(\mathbb{R}^d \rightarrow \mathbb{C})$ it holds that*

$$\mathcal{F}_h(X)(u) = \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} h(v) X(u - v) dv.$$

As a consequence, if $g, \check{g} \in L^1(\mathbb{R}^d \rightarrow \mathbb{C})$ then

$$\mathcal{F}_g(X)(u) = \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} \check{g}(v) X(u - v) dv.$$

Proof. We have that, by definition of impulse response and the spectral representation of X ,

$$\begin{aligned} \int_{\mathbb{R}^d} h(v) X(u - v) dv &= \int_{\mathbb{R}^d} \left(h(v) \int_{\mathbb{R}^d} e^{i\omega(u-v)} dZ(\omega) \right) dv \\ &= \int_{\mathbb{R}^d} \left(\int_{\mathbb{R}^d} h(v) e^{-i\omega v} dv \right) e^{i\omega \cdot u} dZ(\omega) = (2\pi)^d \mathcal{F}_{\hat{h}}(X). \end{aligned}$$

Where in the above we used the stochastic Fubini theorem (Proposition 3). The second part of the proposition follows from Fourier inversion. We have that $\hat{\hat{g}} = g$ so

$$\mathcal{F}_g = \mathcal{F}_{\hat{\hat{g}}} = \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} \check{g}(v) X(u - v) dv.$$

□

The fact that the constant $(2\pi)^d$ pops up is a consequence of the exponent that appears in our definition of the Fourier transform. If one takes the exponential of $-2\pi i \omega x$ instead no constant appears here, however it would appear when we later take derivatives of stationary processes.

In conclusion, the frequency and impulse response function g, h form a Fourier transform pair, and we can thus move between one and the other when sufficient integrability conditions are met. In a way, this makes sense, as g acts on the frequency domain while the h acts on the original domain. Furthermore, heuristically, we have that with the purely formal notation in (12)

$$\mathcal{F}_g(X) = \mathcal{F}^{-1}(g \mathcal{F}(X)) = (2\pi)^{-d} \mathcal{F}^{-1}(g) * X = (2\pi)^{-d} h * X.$$

Where it was used that the (inverse) Fourier transform of the product is the convolution of the (inverse) Fourier transforms. This said, in a way, we still have not answered the most pressing question. Is the result of applying a linear filter to a stationary field another stationary field?

Proposition 5. *Given the output functions $Y_1 := \mathcal{F}_{g_1}(X), Y_2 := \mathcal{F}_{g_2}(x)$ of a linear filtering, we have that their mixed covariance function is*

$$\text{Cov}(Y_1(u), Y_2(v)) = \int_{\mathbb{R}^d} e^{i\omega(u-v)} g_1(\omega) \overline{g_2(\omega)} dF(\omega).$$

As a result, the output $Y = \mathcal{F}_g(X)$ of a linear filter is a stationary field. Furthermore, if X has spectral density $S_X(\omega)$ then Y has spectral density

$$S_Y(\omega) = |g(\omega)|^2 S_X(\omega).$$

Proof. The first part of the proof is a direct application of the isometry for the stochastic integral against $dZ(w)$ and the definition of linear filter. Both the stationarity of $Y = \mathcal{F}_g(X)$ and the calculation of its spectral density follow quickly by setting $g_1 = g_2 = g$. \square

4.2 Differentiation as a linear filter

We now give a couple of examples of linear filters. First, consider $g = 1_A$ where A is some measurable subset of \mathbb{R}^d . Then \mathcal{F}_g corresponds to “filtering” out the frequencies of X outside A . Another example of paramount importance occurs is that, under some integrability conditions, differentiation is itself a linear filter. First, we define in what sense we take derivatives of random fields.

Definition 11. *Given a \mathbb{C}^n valued random field $\{X(u)\}_{u \in \mathbb{R}^d}$, we say that X is mean square differentiable with respect to u_i at u if the following limit exists in $L^2(\Omega \rightarrow \mathbb{C}^n)$*

$$\partial_j X(u) = \lim_{h \rightarrow 0} \frac{X(u + h e_j) - X(u)}{h} \in L^2(\Omega \rightarrow \mathbb{C}) \quad (14)$$

Where $e_j = (0, \dots, \overset{(i)}{1}, \dots, 0)$ is the i -th basis vector in \mathbb{R}^d . Given some $m \in \mathbb{N}$ and $\beta \in \{1, \dots, d\}^m$, if the iterated limits below exist, we write

$$\frac{\partial X}{\partial u^\beta}(u) := \partial_{\beta_1} \dots \partial_{\beta_d} X(u) \in L^2(\Omega \rightarrow \mathbb{C}).$$

Now given $\alpha = (\alpha_1, \dots, \alpha_d) \in \mathbb{N}^d$ we say that X is α times mean square differentiable at u if for all permutations σ of $\{1, \dots, \sum_{n=1}^d \alpha_n\}$ it holds that

$$D^\alpha X(u) := \frac{\partial X}{\partial u^\beta}(u) = \frac{\partial X}{\partial u^{\beta_\sigma}}(u); \quad \beta := (1, \dots, \overset{(\alpha_1)}{1}, \dots, \overset{(\alpha_2)}{2}, \dots, \overset{(\alpha_d)}{d}) \quad (15)$$

Finally, we say that X is mean square differentiable with respect to u_i and that X is α times mean square differentiable if respectively (14),(15) hold for all $u \in \mathbb{R}^d$.

Admittedly, the notation in the above definition is a bit messy. In words, we say that X is (mean square differentiable) with respect to u_i if the usual limit exists in $L^2(\Omega \rightarrow \mathbb{C})$ and that the α -th derivative exists if so do all the other derivatives of lower order and the order of differentiation does not matter.

What does all this have to do with linear filters? Well, as in calculus, we have that mean square differentiation is linear and commutes with translation. As a result, we may suspect that differentiation is given by some linear filter \mathcal{F}_g . Under some integrability assumptions, this is exactly what occurs, as we show in the next proposition

Theorem 6 (Differentiation of fields). *Given a mean 0 stationary field $X : \mathbb{R}^d \rightarrow \mathbb{C}$ with spectral distribution dF the following are equivalent*

1. X is α times mean square differentiable.
2. r is 2α times square differentiable in some interval of 0.
3. r is 2α times square differentiable and $D^\alpha X$ has covariance function

$$r_{D^\alpha X} = (-1)^{|\alpha|} D^{2\alpha} r.$$

4. $\omega^\alpha := \omega^{\alpha_1} \dots \omega_d^{\alpha_d} \in L^2(\mathbb{R}^d \rightarrow \mathbb{C}, dF)$.
5. X is α times mean square differentiable with

$$D^\alpha X(u) = \int_{\mathbb{R}^d} (i\omega)^\alpha e^{i\omega \cdot u} dZ(\omega). \quad (16)$$

Proof. To prove the equivalence between the points, it is sufficient to prove it when $|\alpha| = 1$ and then apply an iteration. Let us take then

$\alpha = (0, \dots, \overset{(i)}{1}, \dots, 0)$. We have that, by the Loeve condition (link) for L^2 convergence, the derivative exists if and only if the limit

$$\begin{aligned} \lim_{h,k \rightarrow 0} \mathbb{E} \left(\frac{X(u + he_j) - X(u)}{h} \cdot \overline{\frac{X(u + ke_j) - X(u)}{k}} \right) \\ = \lim_{h,k \rightarrow 0} \frac{r((h-k)e_j) - r(he_j) - r(-ke_j) + r(0)}{kh} \end{aligned}$$

exists as an element of $L^2(\Omega \rightarrow \mathbb{C})$. Furthermore, if r is twice differentiable in a neighborhood, we recognize the fraction in the above limit as an approximation to $-\partial_j^2 r(0)$. This shows that 2 implies 1 (for more details see [3] page 35.) We now show that 1 and 4 are equivalent. To see that 1 implies 4 we can use that

$$\begin{aligned} -\partial_j^2 r(0) &= \lim_{h \rightarrow 0} \mathbb{E} \left[\left| \frac{X(u + he_j) - X(u)}{h} \right|^2 \right] = \lim_{h \rightarrow 0} 2 \frac{r(0) - \operatorname{Re}(r(he_j))}{h^2} \\ &= \lim_{h \rightarrow 0} \int_{\mathbb{R}^d} 2 \frac{1 - \cos(\omega_j h)}{(\omega_j h)^2} \omega_j^2 dF(\omega) \geq \int_{\mathbb{R}^d} \omega_j^2 dF(\omega). \end{aligned}$$

Where the first equality was already shown to be implied by point 1 (link) and the inequality is due to Fatou's lemma. The fact that 4 implies 2, and thus also 1, is proved similarly by the dominated convergence theorem. We now prove the equivalence of 4 and 5. Suppose that 4 holds. Then we have that, by the dominated convergence theorem

$$i\omega_j = \lim_{h \rightarrow 0} \frac{e^{i\omega_j h} - 1}{h} \in L^2(\mathbb{R}^d \rightarrow \mathbb{C}, dF).$$

By the isometry verified by dZ and the spectral representation in (3) (the spectral representation of X) we deduce that

$$\begin{aligned} \partial_j X(u) &= \lim_{h \rightarrow 0} \frac{X(u + he_j) - X(u)}{h} = \lim_{h \rightarrow 0} \int_{\mathbb{R}^d} e^{i\omega \cdot u} \frac{e^{i\omega_j h} - 1}{h} dZ(\omega) \\ &= \int_{\mathbb{R}^d} i\omega_j e^{i\omega \cdot u} dZ(\omega) \in L^2(\Omega \rightarrow \mathbb{C}). \end{aligned}$$

On the other hand, we have that

$$i\omega_j e^{i\omega \cdot u} \in L^2(\mathbb{R}^d \rightarrow \mathbb{C}, dF) \iff \omega_j \in L^2(\mathbb{R}^d \rightarrow \mathbb{C}, dF).$$

Where the above are functions of ω . This shows that 5 implies 4.

Finally, trivially we have that 3 implies 1 so it only remains to show that all (or any) of the equivalent points 1, 2, 4, 5 implies 3. This follows by a differentiation under the integral sign in

$$r(u) = \int_{\mathbb{R}^d} e^{i\omega \cdot u} dF(\omega).$$

Which can be justified by point 4. □

Observation 2. *It poses no difficulty to extend the above result to vector-valued stationary fields $X : \mathbb{R}^d \rightarrow \mathbb{C}^n$. We have that X is α times differentiable if and only if its components are differentiable. Thus, for example, X will be α times differentiable with derivative*

$$D^\alpha X_j(u) = \int_{\mathbb{R}^d} \omega^\alpha e^{i\omega \cdot u} dZ_j(\omega),$$

if and only if r_{jj} are $2|\alpha|$ times differentiable, if and only if ω^α is in $L^2(\mathbb{R}^d \rightarrow \mathbb{C}, dF_{jj})$. Where the above is for $j = 1, \dots, n$.

The just proved Theorem 6 shows that, if the derivative of a stationary field exists, it is indeed given as the output of a linear filter \mathcal{F}_g where in this case $g(\omega) = (i\omega)^\alpha$. In particular, we have that by the previous results on linear filters, the derivative of a stationary field is itself stationary. Furthermore, as a simple consequence, we obtain the following corollary, which gives the spectral density of the derivative.

Corollary 2. *Given a stationary field $X : \mathbb{R}^d \rightarrow \mathbb{C}$ with spectral density $S(\omega)$ and $\alpha \in \mathbb{N}^d$ we have that $D^\alpha f$ exists if and only if*

$$S_\alpha(\omega) := |\omega^{2\alpha}| f(\omega) \in L^1(\mathbb{R}^d \rightarrow \mathbb{C}).$$

In this case the spectral density of $D^\alpha f$ is $S_\alpha(\omega)$.

4.3 Differential equations

Not only is there a link between linear filters and differentiation, but also between filters and solutions to ODEs, SDEs, and even SPDEs. The basic idea is that solutions to differential equations are often given by a convolution against an impulse response function h , which is determined by the structure of the equations.

4.3.1 Ordinary differential equations

Let us first consider the ordinary differential equation

$$a_0 x^{(p)}(t) + a_1 x^{(p-1)}(t) + \dots + a_p x(t) = y(t) \tag{17}$$

$$x^{(j)}(0) = x_{j,0}, \quad j = 1, \dots, p \tag{18}$$

Where $y : \mathbb{R} \rightarrow \mathbb{R}$ and $a_0, \dots, a_n \in \mathbb{R}$ are given whereas $x : \mathbb{R} \rightarrow \mathbb{R}$ is unknown. Then, by a standard trick we consider higher order derivatives to be new variables by taking

$$\mathbf{x} = (x, x', \dots, y^{(p)}); \quad \mathbf{x}_0 := (x_{1,0}, \dots, x_{p,0}).$$

In this way we obtain that equation (17) is equivalent to the system of equations

$$\mathbf{x}' = A\mathbf{x} + \mathbf{y}; \quad \mathbf{x}(0) = \mathbf{x}_0.$$

Where

$$A := \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & 1 \\ -\frac{a_p}{a_0} & -\frac{a_{p-1}}{a_0} & -\frac{a_{p-2}}{a_0} & \cdots & -\frac{a_1}{a_0} \end{pmatrix}; \quad \mathbf{y}(t) := \begin{pmatrix} 0 \\ \vdots \\ 0 \\ y(t) \end{pmatrix} \quad (19)$$

By the basic theory of linear systems of ODEs (Duhamel's formula) this system of equations has as its solution

$$\mathbf{x}(t) = e^{At}\mathbf{x}_0 + \int_0^t e^{A(t-s)}y(s)ds.$$

If x, y are random variables X, Y , we just interpret that the equation must hold for almost every path and obtain analogously that for $\mathbf{Y}(t) := (0, \dots, 0, Y(t))$ and initial data $\mathbf{Y}(0) = (Y_{1,0}, \dots, Y_{p,0})$ the differential equation is solved by

$$\mathbf{X}(t) = e^{At}\mathbf{Y}(0) + \int_0^t e^{A(t-s)}\mathbf{Y}(s)ds.$$

In general, the above solution is not stationary, as the integration is performed over the range $[0, t]$ as opposed to the whole of \mathbb{R} . However, if all the eigenvalues of A have negative real part, then the equation will be stable (that is, it will converge to 0 exponentially fast when t goes to infinity) and we can consider as initial data $\mathbf{X}(0) = \int_{-\infty}^0 e^{-As}\mathbf{X}(s)ds$, which leads to a solution

$$\mathbf{X}(t) = \int_{-\infty}^t e^{A(t-s)}\mathbf{Y}(s)ds = \int_{\mathbb{R}} e^{A(t-s)}1_{[0,\infty)}(t-s)\mathbf{Y}(s)ds$$

Thus, we obtain a convolution of the form studied previously. Where now $\mathbf{h}(s) = e^{sA}1_{(0,\infty)}(s)$. This gives the following result:

Proposition 6 (Stationary solutions to ODEs). *Consider a mean zero stationary field $Y : \mathbb{R} \rightarrow \mathbb{C}$ and a family of linear constant coefficient ODEs*

$$a_0 X^{(p)}(t) + a_1 X^{(p-1)}(t) + \dots + a_p X(t) = Y(t).$$

Suppose that the matrix A in (19) has all negative eigenvalues. Then the above equation has a stationary solution, given by

$$X(t) = \int_{\mathbb{R}} h(t-s) Y(s) ds.$$

Furthermore, if the spectral density $S_X(\omega), S_Y(\omega)$ of X, Y exist, then

$$S_X(\omega) = \frac{S_Y(\omega)^2}{|P(i\omega)|^2}. \quad (20)$$

Where $P(\omega) := a_0 \omega^p + \dots a_1 \omega + a_p$.

Proof. The proof is a consequence of the previous discussion and the fact that the solution to the family of ODEs is given by

$$X(t) = \mathbf{X}_1(t) = \left(\int_{-\infty}^t \mathbf{h}(t-s) \mathbf{Y}(s) ds \right)_1 = \int_{\mathbb{R}} h(t-s) Y(s) ds.$$

Where $h(s) = (e^{sA})_{1,p} 1_{[0,\infty)}(s)$.

Finally, the formula (20) for the spectral density of X follows from Corollary 2, where we calculated the spectrum of the derivative of a field. □

4.3.2 Stochastic differential equations

An analogous idea holds for SDEs, given an SDE of the form

$$a_0 X^{(p)}(t) + a_1 X^{(p-1)}(t) + \dots + a_p X(t) = dW(t)$$

We will obtain a solution of the form

$$X(t) = \int_{-\infty}^t h(t-s) dW(s).$$

However, we note that the white noise dW is not a stationary field in the sense previously defined. However, as we will see, a very useful interpretation is that dW is a generalized stationary field with spectral density 1 (different conventions in the definition of the Fourier transform to make it so different authors take the spectral density of white noise to be $1/2\pi$) and covariance function its “Fourier transform” $\delta_0(\omega)$.

Definition 12. Let dW be a White noise on $L^2(\mathbb{R}^d)$ and $h \in L^2(\mathbb{R}^d)$. Then we define the convolution of h against dW as

$$h \star dW(u) := \int_{\mathbb{R}} h(u-v) dW(v).$$

We can calculate the Fourier transform of the convolution as a multiplication of the Fourier transforms scaled by some constant.

Lemma 5. Given $f, g \in L^2(\mathbb{R}^d \rightarrow \mathbb{C}^n)$, it holds that

$$\widehat{f * g} = (2\pi)^d \hat{f} \hat{g}.$$

By applying the definition of the Fourier transform in (1), Fubini, and a change of variables we get that

Proof.

$$\begin{aligned} (2\pi)^d \widehat{f * g}(\omega) &= \int_{\mathbb{R}^d} f * g(u) e^{-i\omega \cdot u} du = \int_{\mathbb{R}^d} \left(\int_{\mathbb{R}^d} f(v) g(u-v) dv \right) e^{-i\omega \cdot u} du \\ &= \int_{\mathbb{R}^d} f(v) e^{-i\omega \cdot v} \left(\int_{\mathbb{R}^d} g(u-v) e^{-i\omega \cdot (u-v)} du \right) dv \\ &= \int_{\mathbb{R}^d} f(v) e^{-i\omega \cdot v} \left(\int_{\mathbb{R}^d} g(u) e^{-i\omega \cdot u} du \right) dv = (2\pi)^{2d} \hat{f} \hat{g}. \end{aligned}$$

As desired. □

We now prove that convolution against white noise defines a stationary process.

Lemma 6. Let dW be a White noise on $L^2(\mathbb{R}^d)$ and $h \in L^2(\mathbb{R}^d)$. Then

$$X(u) := h \star dW(u)$$

is a stationary field $X : \mathbb{R}^d \rightarrow \mathbb{R}$ with covariance function

$$r(u) = \int_{\mathbb{R}^d} h(u-v) h(-v) dv.$$

Furthermore, if also $h \in L^1(\mathbb{R}^d)$ then $r \in L^1(\mathbb{R}^d)$ has spectral density

$$S = (2\pi)^d \left| \hat{h} \right|^2.$$

Proof. We have the first equality holds by Itô's isometry, and we can write

$$r(u) = \int_{\mathbb{R}^d} h(u-v)h(-v)dv = h * \tilde{h}(u).$$

Where $\tilde{h}(u) := h(-u)$. In consequence, by the preceding Lemma, if $h \in L^1(\mathbb{R}^d)$ we have that $r \in L^1(\mathbb{R}^d)$ and X has spectral density

$$\hat{r} = (2\pi)^d \hat{h} \hat{h} = (2\pi)^d \hat{h} \overline{\hat{h}} = (2\pi)^d |\hat{h}|^2.$$

This completes the proof. \square

Using this lemma, we can show a similar result holds in the case of SDEs as for ODEs.

Theorem 7 (Stationary solutions to SDEs). *Consider an SDE*

$$a_0 X^{(p)}(t) + \dots a_p X(t) = dW(t).$$

Where the matrix A as defined in (19) has all negative eigenvalues. Then, there exists a stationary solution of the form

$$X(t) = \int_{-\infty}^t h(t-s)dW(s).$$

Furthermore, if its spectral density $S(\omega)$ exists, then

$$S(\omega) = |g(\omega)|^2 = \frac{2\pi}{|P(i\omega)|^2}.$$

Where $P(\omega) := a_0\omega^p + \dots a_1\omega^{p-1} + a_p$.

Proof. To show that there exists a solution of the given form, we can use the theory for solutions to linear SDEs

$$d\mathbf{X} = A\mathbf{X} + d\mathbf{W}.$$

We obtain the analog of Duhamel's formula for a linear system

$$\mathbf{X}(t) = e^{At}\mathbf{X}_0 + \int_0^t e^{A(t-s)}d\mathbf{W}(s).$$

For a proof based on Itô's formula see Øksendal's book on SDEs [9] page 65. To now obtain X of the desired form, we set as in the theorem for ODEs

$$X(t) = [\mathbf{X}(t)]_1; \quad \mathbf{X}_0 := (0, \dots, X_0); \quad X_0 := \int_{-\infty}^0 e^{-As} dW(s).$$

Where X_0 is well-defined by the assumption that the SDE is stable. Finally, the resulting spectral density follows formally from Propositions 5 and 2 together with the fact that the “spectral density” of one dimensional white noise is 2π (we'll see more of this later) (link). Alternatively, one can also compute $|\hat{h}|^2$ where $h(s) = (e^{sA})_{1,p} 1_{(0,\infty)}(s)$. \square

4.3.3 Stochastic partial differential equations

In fact the theory extends all the way to SPDES, we state a necessarily imprecise version of this generalization.

Theorem 8 (Stationary solutions to SPDEs). *Consider an SPDE of the form*

$$\mathcal{L}X(u) = dW(u) \tag{21}$$

Where \mathcal{L} is a linear operator which does not depend on u . Let $P : \mathbb{R}^d \rightarrow \mathbb{C}$ solve

$$\mathcal{L}e^{i\|\omega\|^2} = P(\omega)e^{i\|\omega\|^2}.$$

Then, if $A^{-1} \in L^2(\mathbb{R}^d)$. there exists a stationary solution of the form

$$X(s) = h_{\mathcal{L}} * dW(s) := \int_{\mathbb{R}^d} h_{\mathcal{L}}(s - u) dW(u).$$

And X has spectral density

$$S = |\widehat{h_{\mathcal{L}}}|^2 = \frac{2\pi}{|P|^2}.$$

The theorem is shown in Section 3 of Whittle's paper. [10]. To show an example of the theorem in use, we have that if \mathcal{L} is a finite sum of partial differential operators

$$\mathcal{L} = \sum_{\alpha \in B} \lambda_{\alpha} \left(\frac{\partial}{\partial u_1} \right)^{\alpha_1} \cdots \left(\frac{\partial}{\partial u_d} \right)^{\alpha_d}, \quad \lambda_{\alpha} \in \mathbb{C}, \quad B \subset \mathbb{N}^d.$$

Then we have that

$$\mathcal{L}e^{i\|\omega\|^2} = \left(\sum_{\alpha} \lambda_{\alpha} (i\omega)^{\alpha} \right) e^{i\|\omega\|^2}.$$

That is P is a polynomial and by applying the previous theorem we obtain that the solution to $\mathcal{L}X = dW$ has spectrum

$$S = \frac{1}{|\sum_{\alpha \in B} \lambda_{\alpha} (i\omega)^{\alpha}|^2}.$$

Thus, we see that in this case, P as in Theorem 8 is a polynomial. We will call P the polynomial associated to the operator \mathcal{L} .

4.4 Simulations

In this subsection, we look at how the previous theoretical results can be used to extract fields with some desired properties. We aim to calculate their spectral densities, covariance functions, and simulate from them.

As was mentioned in the introduction, one of the fields most used in statistics are Gaussian fields with Matérn covariance function. These are a family of stationary Gaussian fields on \mathbb{R}^d and that solve the SPDE given by the operator

$$\mathcal{L}_M := (-\Delta + \kappa^2)^{\alpha/2}. \quad (22)$$

Where $\kappa > 0, \alpha > d/2$ are two real parameters.

Observation 3 (Meaning of the fractional derivative). *Though we will not employ these details, we mention for completeness that the linear differential operator in (22) is defined by its effect on the Fourier transform by*

$$\widehat{\mathcal{L}_M f}(\omega) := (-\|\omega\|^2 + \kappa^2)^{\alpha/2} \mathcal{F}(f), \quad f \in \mathcal{S}(\mathbb{R}^d).$$

Where, we used \mathcal{F} to denote the Fourier transform and $\mathcal{S}(\mathbb{R}^d)$ is the Schwartz space on \mathbb{R}^d . That is, the space of infinitely differentiable functions whose derivatives decay infinitely fast. In fact, by a calculation, \mathcal{L}_M can be shown to extend continuously to a continuous linear operator

$$\mathcal{L}_M : H^{s+\alpha}(\mathbb{R}^d) \rightarrow H^s(\mathbb{R}^d).$$

Where $H^s(\mathbb{R}^d)$ is the Sobolev space of s -times differentiable functions

$$H^s(\mathbb{R}^d) := \{f \in L^2(\mathbb{R}^d) : \|f\|_{H^s(\mathbb{R}^d)} := (1 + \|\omega\|^2)^{s/2} \hat{f}(\omega) \in L^2(\mathbb{R}^d \rightarrow \mathbb{C})\}.$$

See for example [11] page 80 for more details.

Next, we wish to consider a stationary field whose spectral density is largest on a sphere of radius $r \in \mathbb{R}_+$. We can obtain such a field as a solution to an SPDE whose associated polynomial takes its minimum, we can consider for example the polynomial

$$P_S(\omega) := (\|\omega\|^2 - r)^2 + \lambda_1.$$

Which proceeds from the operator

$$\mathcal{L}_S := (-\Delta - r)^2 + \lambda_1 \quad (23)$$

Where $\lambda_1 \in \mathbb{R} \setminus \{0\}$ is fixed. Finally, we will also consider a stationary field whose spectrum has two peaks, one at v and another $-v$, for some fixed $v \in \mathbb{R}^d$. To do so we can consider, for example,

$$P_T(\omega) := \|\omega - v\|^2 \|\omega + v\|^2 + \lambda_2.$$

Where $\lambda_2 \in \mathbb{R} \setminus \{0\}$. By Theorem 8, the operator associated is

$$\mathcal{L}_T = (-\Delta + \|v\|^2)^2 + 4\nabla \cdot vv^T \nabla + \lambda_2 \quad (24)$$

Let us denote by S_M, r_M, X_M the spectral density, the covariance function, and the stationary field that solve the linear equation given by \mathcal{L}_M . We also use analogous notation to define S_S, r_S, X_S and S_T, r_T, X_T . We define $n = (n_1, \dots, n_d) \in \mathbb{N}^d, N = n_1 \cdots n_d$ and for $x, y \in \mathbb{R}^d$

$$\frac{1}{x} := \left(\frac{1}{x_1}, \dots, \frac{1}{x_d} \right); \quad xy = (x_1 y_1, \dots, x_d y_d).$$

We will also consider

$$\omega_k := \frac{2\pi k}{hn}, \quad k \in \mathcal{K}_0 := \left[\frac{n}{2}, \frac{n}{2} - 1 \right] \cap \mathbb{Z}^d \quad (25)$$

By Theorem 8 and Bochner's theorem, we know that, for $i \in \{M, S, T\}$

$$S_i = \frac{2\pi}{P^2}, \quad r_i = \hat{S}_i.$$

Thus, we need no numerical approximation to obtain S_i . To obtain, r_i however, we need to approximate the Fourier integral over a grid. Doing so naively using Riemann sums has a cost $O(N^2)$. Thus, we will use a fast Fourier transform algorithm. This uses the discrete Fourier transform to reduce the order of complexity to $O(N \log(N))$.

Finally, we wish to simulate from the three fields X_M, X_S, X_T . To do so, we will use the spectral process of Theorem 2 (**Make sure your field has mean zero otherwise see Observation 1**). Let us write X for any one of these fields we have that, by this theorem

$$X(u) = \int_{\mathbb{R}^d} e^{i\omega \cdot u} dZ(\omega). \quad (26)$$

Furthermore, since X is a stochastic integral of a deterministic function against White noise, X is a Gaussian variable. In consequence, by Corollary (1), we deduce that $dZ(f)$ is Gaussian for all $f \in L^2(\mathbb{R}^d)$. By our observation in (7), we can approximate the stochastic integral in (26) above as

$$X(u) \sim \sum_{k \in \mathcal{K}_0} e^{i\omega_k \cdot u} Z_k.$$

Where $Z_k := dZ((\omega_k, \omega_{k+1}])$ is a Gaussian random variable with variance

$$\|Z_k\|_{L^2(\Omega \rightarrow \mathbb{C})}^2 = \|Z_k\|_{L^2(\mathbb{R}^d \rightarrow \mathbb{C}, dF)}^2 = \int_{(w_k, w_{k+1}]} S(\omega) d\omega.$$

This is what we talked about today, I don't know why Finn and I never used this though. Let us denote the variance of dZ_k by σ_k , so that,

$$dZ_k = \sigma_k(x_k + iy_k).$$

Where x_k, y_k are iid unit, real-valued normal random variables for $k \in \mathcal{K}_+$. By the isometry in (4), we have that Z_k, Z_j are independent for different $k \neq j$. However, why are the real and imaginary part of Z_k iid and had the same variance. Is this a property or just something we assume? Why make such an assumption? Also I checked his spectral notes, and at the end of page three he says the complex and imaginary parts are not independent?? Since we know that X is real valued, its imaginary part must be zero, so as

n goes to infinity (in all directions)

$$\begin{aligned} X(u) &\simeq \sum_{k \in \mathcal{K}_0} e^{i\omega_k \cdot u} Z_k = \sum_{k \in \mathcal{K}_0} \sigma_k (\cos(\omega_k \cdot u) + i \sin(\omega_k \cdot u))(x_k + iy_k) \\ &= \sum_{k \in \mathcal{K}_0} \sigma_k (x_k \cos(\omega_k \cdot u) - y_k \sin(\omega_k \cdot u)) \end{aligned} \quad (27)$$

Using this together with the independence of $x_k, y_k \sim \mathcal{N}(0, 1)$ gives that

$$\begin{aligned} r(u - v) &= \text{Cov}(X(u), X(v)) \\ &\simeq \sum_{k \in \mathcal{K}_0} \sigma_k^2 (\cos(\omega_k \cdot u) \cos(\omega_k \cdot v) + \sin(\omega_k \cdot u) \sin(\omega_k \cdot v)) \\ &= \sum_{k \in \mathcal{K}_0} \sigma_k^2 \cos(\omega_k \cdot (u - v)). \end{aligned} \quad (28)$$

Where we used the formula for the cosine of the difference in the last equality. By a change of variables and the period of the complex exponential, we have that if u is divisible by h (i.e. $u = jh$ for some $j \in \mathbb{Z}^d$), then by Bochner's theorem and the periodicity of the complex exponential

$$\begin{aligned} r(u) &= \sum_{k \in \mathbb{Z}^d} \int_{[-\frac{\pi+2\pi k}{h}, \frac{\pi+2\pi k}{h})} e^{iu \cdot \omega} S(\omega) d\omega = \int_{[-\frac{\pi}{h}, \frac{\pi}{h})} e^{iu \cdot \omega} \sum_{k \in \mathbb{Z}^d} S\left(\omega + \frac{2\pi k}{h}\right) d\omega \\ &= \int_{[-\frac{\pi}{h}, \frac{\pi}{h})} e^{iu \cdot \omega} \tilde{S}_h(\omega) d\omega. \end{aligned}$$

Where we defined \tilde{S}_h as the sum appearing in the above integral. Since r must be real, S (and thus \tilde{S}_h) is real, we deduce that, if $u - v$ is divisible by h , then

$$\begin{aligned} r(u - v) &= \int_{[-\frac{\pi}{h}, \frac{\pi}{h})} \text{Re}(e^{i(u-v) \cdot \omega}) \tilde{S}_h(\omega) d\omega = \int_{[-\frac{\pi}{h}, \frac{\pi}{h})} \cos(\omega \cdot (u - v)) \tilde{S}_h(\omega) d\omega \\ &\simeq \sum_{k \in \mathcal{K}_0} \left| \frac{2\pi}{hn} \right| \tilde{S}_h(\omega_k) \cos(\omega_k \cdot (u - v)). \end{aligned} \quad (29)$$

Where in the last line we did an approximation to the integral over the interval $[-\pi/h, \pi/h)$ by subdividing it into $n/2$ parts. Due to (28) and (29), and by the definition of ω_k in (25), we set

$$\sigma_k^2 := \left| \frac{2\pi}{hn} \right| \tilde{S}_h\left(\frac{2\pi k}{hn}\right), \quad \forall k \in \mathcal{K}_0.$$

We can now calculate $X(s)$ according to the discrete approximation to the spectral representation in (27). Thus, we are now able to calculate all of the terms we desire. Now, I know what you're thinking, but what about generalized fields? Can they be stationary, what does the spectral theory say for them? Ok, maybe none of you were thinking that, in fact you may now be asking yourselves what a generalized field is. A good question indeed, and one that will be answered in the next post along with your other most burning questions related to Bochner's theorem and the spectral theorem for these fields (if you don't have these questions and want your money back, it's too late. Also, you didn't give me any money anyway, so you're not getting it back.).

5 Appendix

5.1 The Bochner integral

We now define integration for functions valued in a Banach space

$$f : (\Omega, \mathcal{F}, \mu) \rightarrow ((X, \|\cdot\|), \mathcal{B}(X))$$

If we denote the class of simple functions by \mathcal{A} we can define their integral as

$$\int f d\mu = \int \sum_{k=1}^n x_k 1_{A_k} d\mu = \sum_{k=1}^n x_k \mu(A_k)$$

If we take equivalence class and define the norm

$$\|f\|_{\mathcal{A}} := \int_{\Omega} \|f\| d\mu.$$

Then we get that integration is a linear and absolutely continuous function

$$int : (\mathcal{A}, \|\cdot\|_{\mathcal{A}}) \rightarrow (X, \|\cdot\|).$$

as for all $f \in \mathcal{A}$

$$\left\| \int_{\Omega} f \right\| \leq \int_{\Omega} \|f\| d\mu = \|f\|_{\mathcal{A}}$$

This shows that we can extend integration in a unique way to the completion $\bar{\mathcal{A}}$ of $(\mathcal{A}, \|\cdot\|_{\mathcal{A}})$.

Definition 13. We say a function $f : (\Omega, \mathcal{F}) \rightarrow (X, \mathcal{B}(X))$ is strongly measurable if f is measurable $f(\Omega)$ is separable.

Definition 14. For $1 \leq p < \infty$ we define

$$\mathcal{L}^p(\Omega, \mathcal{F}, \mu, X) = \left\{ f : \Omega \rightarrow X : f \text{ is strongly measurable and } \int \|f\|^p d\mu < \infty \right\}.$$

$$\hat{\mathcal{L}}^p(\Omega, \mathcal{F}, \mu, X) = \left\{ f : \Omega \rightarrow X : \int \|f\|^p d\mu < \infty \right\}.$$

We also define the semi-norms

$$\|f\|_{L^p(\Omega \rightarrow X)} := \left(\int \|f\|^p d\mu \right)^{1/p}, \quad f \in \mathcal{L}^p(\Omega, \mathcal{F}, \mu, X).$$

$$\|f\|_{\hat{L}^p(\Omega \rightarrow X)} := \left(\int \|f\|^p d\mu \right)^{1/p}, \quad f \in \hat{\mathcal{L}}^p(\Omega, \mathcal{F}, \mu, X).$$

Finally we take equivalence classes by the above semi-norms to obtain the metric spaces $L^p(\Omega \rightarrow X)$ and $\hat{L}^p(\Omega \rightarrow X)$.

A straightforward adaptation of the proof of the completion of L^p spaces proves that $L^p(\Omega \rightarrow X)$ and $\hat{L}^p(\Omega \rightarrow X)$ are also complete.

Proposition 7. Given a Banach space X , the space of p integrable strongly measurable and measurable functions $L^p(\Omega \rightarrow X)$, $\hat{L}^p(\Omega \rightarrow X)$ are Banach spaces.

Proof. The proof is identical in both cases so we prove it only for $f \in L^p(\Omega \rightarrow X)$ It suffices to show that if f_n is such that

$$\sum_{n=1}^{\infty} \|f_n\|_{L^p(\Omega \rightarrow X)} < \infty.$$

Then there exists $f \in L^p(\Omega \rightarrow X)$ such that

$$f = \sum_{n=1}^{\infty} f_n \in L^p(\Omega \rightarrow X).$$

To do so we apply Minkowski's inequality for real valued functions to show that

$$\sum_{n=1}^{\infty} \|f_n\|_X \in L^p(\Omega \rightarrow \mathbb{R}).$$

and is thus finite almost everywhere. Since X is complete we have that the below sum converges pointwise almost everywhere to some function

$$f(\omega) := \sum_{n=1}^{\infty} f_n(\omega) \in X.$$

Furthermore we have that f is strongly measurable as it is the limit of strongly measurable functions and, by Fatou's lemma for real valued functions and the triangle inequality for norms

$$\begin{aligned} \left\| f - \sum_{n=1}^N f_n \right\|_{L^p(\Omega \rightarrow X)} &= \left\| \sum_{n=N}^{\infty} f_n \right\|_{L^p(\Omega \rightarrow X)} \leq \liminf_{M \rightarrow \infty} \left\| \sum_{n=N}^M f_n \right\|_{L^p(\Omega \rightarrow X)} \\ &\leq \liminf_{M \rightarrow \infty} \sum_{n=N}^M \|f_n\|_{L^p(\Omega \rightarrow X)} = \sum_{n=N}^{\infty} \|f_n\|_{L^p(\Omega \rightarrow X)} \xrightarrow{N \rightarrow \infty} 0. \end{aligned}$$

Which shows convergence in $L^p(\Omega \rightarrow X)$. \square

In the standard construction of the Lebesgue integral it is used that every measurable function to \mathbb{R} can be pointwise approximated by simple functions. One can achieve the same result for arbitrary metric spaces if the image of f is separable.

Lemma 7. *Let (E, d) be a metric space and let f be strongly measurable. Then f is the pointwise limit of simple functions f_n . Furthermore, $d_n(\omega) := d(f_n(\omega), f(\omega))$ is a non-increasing sequence for each $\omega \in \Omega$*

Proof. Consider $\{e_k\}_{k=1}^{\infty}$ to be dense in $f(\Omega)$. Now set

$$f_n(\omega) := \arg \min_{i \in \{1, \dots, n\}} \{d(f(\omega), e_i)\}$$

Since for each fixed ω the above distance goes to 0 we have that

$$\lim_{n \rightarrow \infty} f_n(\omega) = f(\omega), \quad \forall \omega \in \Omega.$$

Furthermore, f_n is simple as $f_n(\Omega) \in \{e_1, \dots, e_n\}$ and the non-increasing property of $d_n(\omega)$ is clear by construction. \square

In the above proof we see that the reason for requiring that the image of our class of integrable functions be separable is so that we can approximate them by simple functions.

Proposition 8. *Every function $f \in L^1(\Omega \rightarrow X)$ is the limit of simple functions.*

Proof. Since $f(\Omega)$ is dense we can apply the above lemma to obtain a sequence of simple functions f_n converging point-wise to f . Furthermore we have that, by the monotone convergence theorem for integrals in \mathbb{R}

$$\lim_{n \rightarrow \infty} \|f_n - f\|_{L^1(\Omega \rightarrow X)} = \lim_{n \rightarrow \infty} \int_{\Omega} \|f_n - f\| d\mu = 0.$$

□

As a corollary we obtain the following

Corollary 3. *\mathcal{A} is a dense subset of $(L^1(\Omega \rightarrow X), \|\cdot\|_{L^1(\Omega \rightarrow X)})$.*

In consequence, by the linearity and continuity of the integral and the above density, we have a unique continuous extension of the integral to $L^1(\Omega \rightarrow X)$.

Definition 15. *We define the integral on $L^1(\Omega \rightarrow X)$ as the unique continuous extension with the norm $\|\cdot\|_{L^1(\Omega \rightarrow X)}$ of the integral on \mathcal{A} . That is, given $f \in L^1(\Omega \rightarrow X)$ we define*

$$\int_{\Omega} f d\mu := \lim_{n \rightarrow \infty} \int_{\Omega} f_n d\mu.$$

Where $f_n \in \mathcal{A}$ is any sequence such that $\|f - f_n\|_{L^1(\Omega \rightarrow X)} \rightarrow 0$.

As a result of the definition the following is immediate.

Corollary 4. *Let $f \in L^1(\Omega, X)$ with X a Banach space, then*

1. $\|\int_{\Omega} f d\mu\| \leq \int_{\Omega} \|f\| d\mu$
2. *Let Y be another Banach space and $L \in L(X, Y)$. Then*

$$\int_{\Omega} L \circ f d\mu = L\left(\int_{\Omega} f d\mu\right).$$

“In this section we define some of the concepts that are needed to properly define an SPDE, we will consider SPDEs only in space. If one wished to consider time as well it would be necessary to construct the stochastic integral. This would require a substantial amount of functional and stochastic analysis. A construction mirroring the one done for SDEs can be found in Chapter 2 of [12]. An alternative approach based on expansions in an orthonormal basis can be found in [11] page 131.

5.2 Reproducing kernel Hilbert spaces

In this subsection we will consider all vector spaces to be complex.

Definition 16. *A reproducing kernel space is a Hilbert space S of complex valued functions on some set S (that is, $H \subset \mathbb{C}^S$ where S is some set) and such that point evaluation ℓ_s ,*

$$\ell_s(f) := f(s), \quad f \in H,$$

is continuous for all $s \in S$.

For example the spaces $L^2(\mathbb{R}^d)$ do not fall into this category as pointwise evaluation is not even well defined. That is, unless we consider $S = L^2(\mathbb{R}^d)$ with pointwise evaluation defined as

$$f(s) := \langle f, s \rangle, \quad \forall s \in S.$$

In this case we would obtain that $L^2(\mathbb{R}^d)$ is a reproducing kernel space over itself with reproducing kernel $K(f, \cdot) = f$. This construction can be used to view any Hilbert space as a reproducing kernel space over its dual. To give another example, by Sobolev embedding theory, $H^s(\mathbb{R}^d)$ are reproducing kernel spaces for $s > \frac{d}{2}$.

Definition 17. *A function $K : S \times S \rightarrow \mathbb{C}$ is said to be a positive-definite Kernel on S if $\{K(t_i, t_j)\}_{i,j=1}^n$ is a Hermitian, positive semi-definite matrix for all $t_i \in S, n \in \mathbb{N}$.*

For example the covariance function $K(t, s) = \text{Cov}(X(t), X(s))$ of a complex valued field is a positive semi-definite kernel.

Proposition 9. *Given a reproducing kernel Hilbert space H there exists a unique positive semi-definite kernel K_H called the reproducing kernel of H such that*

$$f(s) = \langle f, K_H(s, \cdot) \rangle, \quad \forall f \in H.$$

Furthermore, given any positive-definite kernel K there exists a reproducing kernel space H_K such that K is the reproducing kernel of H_K .

Proof. We first prove the existence and uniqueness of K_H , by Riesz's representation theorem, for each $s \in S$ there exists $g_s \in H$ such that

$$f(s) = \langle f, g_s \rangle \quad \forall f \in H.$$

Let us define $K_H(t) := g_s(t)$, the defining property is verified by construction and it remains to see that K_H is a positive definite kernel. This follows from the equality

$$K_H(s, t) = \langle K_H(s, \cdot), K_H(t, \cdot) \rangle \quad (30)$$

To prove the existence of H_k we set

$$H_{K,0} := \text{span}(\{K(s, \cdot) : s \in S\}); \quad \left\langle \sum_{j=0}^n K(s_j, \cdot), \sum_{k=0}^m K(t_k, \cdot) \right\rangle_{H_{K,0}} := \sum_{j=0}^n \sum_{k=0}^m K(s_j, t_k).$$

We note that the above defines a inner product by the symmetry and positivity condition on K and that pointwise evaluation is continuous by construction and Cauchy Schwartz. All that is left is to take the completion $H_K := \overline{H_{K,0}}$ as it is a verification that K is the reproducing kernel of K_H . \square

The equality (30) motivates the name reproducing kernel. Another equality of note, which can be easily derived from the above is that, given an orthonormal basis e_k of H

$$K_H(s, t) = \sum_{n=0}^{\infty} e_n(s) \overline{e_n(t)}.$$

Given any topological vector space $(V, \mathcal{B}(V), \mu)$ whose measure allows for the existence of second moments, we can construct a positive definite kernel on V' as

$$C_\mu(\ell, h) := \int_V \ell(v) \overline{h(v)} \mu(dv) - \int_V \ell(v) \mu(dv) \overline{\int_V h(v) \mu(dv)}.$$

We call this the *covariance kernel* associated with μ . By setting $S = V'$ we may apply the previous Proposition 9 to deduce that there exists a Hilbert space H_μ whose reproducing kernel is C_μ . We call this space the *reproducing kernel Hilbert space of μ* . If μ is the distribution of some V valued random variable X then we also say that H_μ is the *reproducing kernel Hilbert space of X* .

Alternatively, we can also view C_μ as a function from $V' \rightarrow \tilde{V}'$. Where given a vector space E we denote the space of antilinear maps by \tilde{E}' In the case where V is a Hilbert space H we can identify H with both H' and \tilde{H}' .

Through these identifications we deduce for each $h \in H$ the existence of some $h_\mu \in H$ such that $C_\mu(h, \cdot) = \langle h_\mu, \cdot \rangle$. That is

$$C_\mu(h, x) = \langle h_\mu, x \rangle, \quad \forall h, x \in H.$$

By another abuse of notation, using the previous identifications we can consider

$$\begin{aligned} C_\mu : H &\longrightarrow H \\ h &\longmapsto C_\mu(h) = h_\mu. \end{aligned}$$

That is, to gather all the notation together

$$C_\mu(\langle \cdot, h \rangle, \langle \cdot, x \rangle) = C_\mu(h, x) = \langle C_\mu(h), x \rangle \quad \forall h, x \in H.$$

5.3 Gaussian Random variables

Definition 18. *Given a topological vector space V , we say that a Gaussian measure is a measure on $\mathcal{B}(V)$ such that ℓ is a real valued Gaussian variable for all $\ell \in E^*$.*

Definition 19. *A V valued random variable is said to be Gaussian if \mathbb{P}_X is a Gaussian measure on V .*

We note that in the case $V = \mathbb{R}^d$ this definition coincides with the classic one.

Proposition 10. *Let H be a separable Hilbert space and identify H with H' , then:*

1. *Given a Gaussian measure μ on $(H, \mathcal{B}(H))$, its covariance operator is trace class. That is, $C_\mu \in \mathcal{L}_1^+(H)$.*
2. *Given $K \in \mathcal{L}_1^+(H)$ there exists a Gaussian measure μ on $(H, \mathcal{B}(H))$ such that $C_\mu(u, v) = \langle Ku, v \rangle$.*

Proof. We prove each point in order

1. Consider an orthonormal basis e_n of H (which we recall we are identifying with H'), then we have that

$$\sum_{n=0}^{\infty} \langle C_\mu(e_n), e_n \rangle_H = \sum_{n=0}^{\infty} C_\mu(\langle \cdot, e_n \rangle, \langle \cdot, e_n \rangle)_H = \int_H \|v\|^2 \mu(dv) < \infty.$$

Where the last equality is due to Fernique's theorem, which says that Gaussian measure have exponential decrease.

2. By theorem we can diagonalize K to obtain a orthonormal basis e_n with $Ke_n = \lambda e_n$. By theorem we can build a probability measure \mathbb{P} on $(H, \mathcal{B}(H))$ together with a sequence G_n of iid unit Gaussian vectors. Let us set

$$\mu := \mathbb{P}_X; \quad X = \sum_{n=0}^{\infty} \sqrt{\lambda_n} G_n e_n.$$

Then we have that μ is a Gaussian measure and, by the independence and normality of G_n

$$\begin{aligned} C_\mu(u, v) &= \int_H \langle u, h \rangle \overline{\langle v, h \rangle} d\mathbb{P}_X = \sum_{n,m=0}^{\infty} \sqrt{\lambda_n} \sqrt{\lambda_m} \langle u, e_n \rangle \overline{\langle v, e_m \rangle} \mathbb{E}[G_n G_m] \\ &= \sum_{n=0}^{\infty} \lambda_n \langle u, e_n \rangle \overline{\langle v, e_n \rangle} = \langle Ku, v \rangle. \end{aligned}$$

□

We now have a relationship between operators in $\mathcal{L}_1^+(H)$ and Gaussian measures. A natural question is whether this relationship is bijective. To do so we introduce the Fourier transform. For Gaussian measures on separable Banach spaces we have the following

Proposition 11. *Let E be a separable Banach space, then μ is a Gaussian measure on E iff for every linear function $\ell : V \rightarrow \mathbb{R}$*

$$\hat{\mu}(\ell) := \int_E e^{i\ell(x)} d\mu(x) = e^{i\ell(m) - \frac{1}{2}C_\mu(\ell, \ell)}.$$

Where $m := \int_E x \mu(dx)$.

Proof. The implication is an application of theory of 1-dimensional Gaussian random variables. By definition we have that μ is a Gaussian measure if and only if all real valued continuous linear functions $\ell : E \rightarrow \mathbb{R}$ are Gaussian random variables. Which in turn implies that their characteristic function φ_ℓ verifies that

$$\hat{\mu}(\ell) = \int_E e^{i\ell(x)} \mu(dx) = \varphi_\ell(1) = e^{ia - \frac{1}{2}\sigma^2}.$$

Where, by the linearity of the Bochner integral,

$$a = \int_E \ell(x) \mu(dx) = \ell(m); \quad \sigma = \int_E \ell(x)^2 \mu(dx) = C_\mu(\ell, \ell).$$

For the converse see [13] page 9. Where it is proved that the Fourier transform characterises measures on separable Banach spaces. \square

5.4 Operators

Proposition 12. *Let X be a normed space, then the only open subspace of X is itself.*

Proof. Let $U \subset X$ be an open subspace and choose $y \notin U$, then, by linearity, $\lambda_n y \notin U$ for all $\lambda_n \in \mathbb{K}$. Choose $\lambda_n \rightarrow 0$, then $\lambda_n y \rightarrow 0 \in U$. As a result U^c is not closed which is a contradiction. \square

Proposition 13. *Let X, Y be normed spaces, then if $T \in \mathcal{L}(X, Y)$ is open it is also surjective*

Proof. $T(X)$ is open and a linear subspace. \square

Proposition 14 (Linear mapping theorem). *Let E_1, E_2 be Banach spaces, then any surjective mapping $T \in \mathcal{L}(E_1, E_2)$ is open.*

Proposition 15 (Homeomorphism theorem). *Let E_1, E_2 be Banach spaces, then $T \in \mathcal{L}(E_1, E_2)$ is a homeomorphism iff it is bijective.*

Proof. By the open mapping theorem a linear (continuous) operator is open iff it is surjective. Since T is surjective it is open so T^{-1} is continuous. \square

5.4.1 Compact operators

Let X, Y be normed spaces.

Definition 20. *We say that an operator T is compact if it transforms bounded sets into relatively compact sets.*

Note that by definition any compact operator must be continuous as relatively compact sets must be bounded. However the reverse is not true, for example $T = Id$ on l^2 . The reason for this is because l^2 does not have finite dimension however, since every bounded set in \mathbb{R}^n is relatively compact we obtain the following.

Proposition 16. *Let $T : X \rightarrow Y$ have finite range, then T is compact \iff it is continuous.*

Proposition 17. *The identity is relatively compact \iff X has finite dimension*

Proof. $\bar{B}(0, 1)$ is compact \iff the dimension of X is finite. \square

Proposition 18. *T is compact $\iff T(B(0, r))$ is relatively compact for some $r \in \mathbb{R}$ $\iff T(B(0, r))$*

Proof. The implication is clear, let $A \subset X$ be bounded. Then for some $r \in \mathbb{R}$

$$T(A) \subset T(B(0, R)) = \frac{R}{r}T(B(0, r))$$

Since any closed set within a compact set is compact and this concludes the proof. \square

Let us write $\mathcal{K}(X, Y)$ for the space of relatively compact linear operators, then

Proposition 19. *Let $T : X \rightarrow Y$ be a linear operator, then*

- $\mathcal{K}(X, Y)$ is a vector space.
- The composition of a continuous and compact operator is compact.
- An isomorphism is compact $\iff \dim(X) < \infty$

Proof. The first two items can be proven using convergent sub-sequences. The second follows from considering $Id = T \circ T^{-1}$ together with the second item. \square

Proposition 20. *Let $T : X \rightarrow Y$ with Y complete. Then if T_n are compact operators converging in $\mathcal{L}(X \rightarrow Y)$ to T we have that T is relatively compact.*

Proof. This follows from the fact that in complete normed spaces compactness is equivalent to total boundedness. So if we cover $T_N(A)$ by $B(y_n, \epsilon/2)$ then for large N and $x \in A$

$$\|T(x) - y_n\| \leq \|T - T_N\| \|x\| + \|T_N(x) - y_n\| < \epsilon$$

\square

Proposition 21. *Let T be a compact operator between normed spaces, then $T(X)$ is separable*

Proof. We have that $T(X)$ is the union of $T(B_n(0))$ which are separable as they are relatively compact and thus totally bounded. \square

5.4.2 Fredholm alternative

Proposition 22. *Let $T \in \mathcal{K}(E)$ with E a Banach space, then $\ker(T - \lambda I)$ is a finite dimensional vector space for all $\lambda \in \mathbb{K}$.*

Proof. Since T is compact so is $\frac{T}{\lambda}$, also when restricted to $\ker(T/\lambda - I) = \ker(T - \lambda I)$. But I is the identity here so this space must have finite dimension \square

5.4.3 Invertible operators

Let E be a Banach space we recall that $\mathcal{L}(E)$ is a ring and that we say that $T \in \mathcal{L}(E)$ is invertible if $T^{-1} \in \mathcal{L}(E)$. Our next proposition states that, if an operator is not too far from the identity, then it is invertible.

Proposition 23. *If $T \in \mathcal{L}(E)$ is such that $\|T\| < 1$ then $I - T$ is invertible with inverse*

$$(I - T)^{-1} = \sum_{n=1}^{\infty} T^n$$

Proof. This follows from the completeness of $\mathcal{L}(E)$ and the fact that the above series is normally convergent. \square

Using this we can prove that operators close to any invertible operator are invertible, where the closeness depends on how large $\|T^{-1}\|$ is.

Corollary 5. *The open operators are an open set of $\mathcal{L}(E)$ for a Banach space E .*

Proof. Let $S \in \mathcal{L}(X)$ with $\|T - S\| < 1/\|T^{-1}\|$, then it holds that

$$\|ST^{-1} - Id\| = \|(S - T)T^{-1}\| \leq \|S - T\|\|T^{-1}\| < 1$$

This proves that ST^{-1} is invertible and thus, since T is invertible so is S . \square

5.4.4 Specter of a linear operator

Definition 21. *Let $T \in \mathcal{L}(E)$ with E a Banach space, then we say that the specter of T is*

$$\sigma(T) := \{\lambda \in \mathbb{K} : T - \lambda I \text{ is not invertible}\}$$

Definition 22. Let $T \in \mathcal{L}(E)$ with E a Banach space, then we say that the eigenvalues of T are the λ such that $T - \lambda I$ is not injective. We call the set of eigenvalues the pointwise specter of T

$$\sigma(T) := \{\lambda \in \mathbb{K} : T - \lambda I \text{ is not invertible}\}$$

For example the spectrum of $T = Id$ is $\sigma(T) = \{1\}$. Note that, by definition, the pointwise spectrum is within the spectrum.

Proposition 24. Let $T \in \mathcal{L}(E)$, then $\sigma(T)$ is a closed subset in $\bar{B}(0, \|T\|)$

Proof. Let $\lambda \notin \sigma(T)$, then $T - \lambda I$ is invertible so, since invertible operators are open set, so is $T - (\lambda + \mu)I$ for small μ . This shows that $\sigma(T)$ is closed. Furthermore, if $\lambda > \|T\|$ then

$$T - \lambda I = \lambda \left(\frac{T}{\lambda} - I \right).$$

Since $\|T/\lambda\| < 1$ this is invertible. □

Definition 23. The spectral radius of T is defined as

$$\rho(T) := \sup\{|\lambda| : \lambda \in \sigma(T)\}$$

Proposition 25. Let $T : E_1 \rightarrow E_2$ be an operator between two Banach spaces such that

$$\|Tx\| \geq c\|x\| \quad \forall x \in E_1.$$

Then $T : X \rightarrow Im(T)$ is a homeomorphism.

Proof. T is injective by the condition, it's image is closed as if $T(x_n)$ converges so does x_n by the condition. As a result, $T : E_1 \rightarrow Im(T)$ is a bijective operator between two Banach spaces so it is invertible by the homeomorphism theorem. □

Proposition 26. Let T be a non-zero compact operator on a Banach space, then nonzero elements $\lambda \in \sigma(T)$ are also eigenvectors with finite dimensional eigenspace $ker(T - \lambda I)$.

Proof. By hypothesis $T - \lambda I$ is compact. If λ is not an eigenvalue then $T - \lambda I$ is injective, so equivalently, by the Fredholm alternative it is surjective and thus, by the homeomorphism theorem, invertible. This contradicts that λ is in the spectrum □

5.5 Self-adjoint operators

Proposition 27. *Let $T \in \mathcal{L}(H)$ with H a Hilbert space, then*

$$\ker(T) = (\operatorname{Im} T^*)^\perp$$

Proof. Let $e \in \ker(T)$, then $\langle e, T^*v \rangle = 0$. The other inclusion is due to $H^\perp = 0$. \square

Proposition 28. *Let T be a self-adjoint operator on a Hilbert space, then its spectrum is non-empty (and contains $\|T\|$ or $-\|T\|$)*

Proposition 29. *Let $T \in \mathcal{L}(H)$ be compact and self-adjoint, then if $D = \sigma_P \setminus \{0\}$ is the set of non-zero eigenvalues we have that*

$$H = \ker(T) \oplus \overline{\bigoplus_{\lambda \in D} \ker(T - \lambda I)}$$

Where $\ker(T - \lambda I)$ are finite dimensional spaces orthogonal amongst each other.

Proof. Firstly we have that, since T is compact, $\ker(T - \lambda I)$ is finite dimensional, and since it is self-adjoint they are orthogonal. Write E for the term on rhs, by self-adjointness also

$$\ker T \subset (\ker(T - \lambda I))^\perp$$

for all $\lambda \in D$ so $\ker T \subset E^\perp$. To see the other inclusion we have that $iT|_{E^\perp}$ is compact and self adjoint, so if it is non-zero it has some nonzero eigenvalue ($\|T\|$ or $-\|T\|$) but all the non-zero eigenvalues are in E not in E^\perp . So $E^\perp \subset \ker T$. As a result $E^\perp = \ker T$ so, since E is a closed vector space,

$$H = E^\perp \oplus E^{\perp\perp} = \ker(T) \oplus E$$

\square

Theorem 9 (Hilbert Schmidt). *Let $T \in \mathcal{L}(H)$ be a compact, self adjoint operator, then it diagonalizes. That is, there exists an orthonormal basis e_i of eigenvectors with eigenvalues $\lambda_i \in \mathbb{R}$. Furthermore,*

$$Tx = \sum_{i \in I} \lambda_i \langle x, e_i \rangle e_i \quad \forall x \in H.$$

Proof. We know that if σ_P is the set of eigenvalues, then

$$H = \overline{\bigoplus_{\lambda \in \sigma_P} \ker(T - \lambda I)}$$

where all the appearing spaces are orthogonal. It suffices to choose an orthonormal basis in each of $\ker(T - \lambda I)$. \square

In the finite dimensional case we have that if T is self adjoint and non-negative then there exists a diagonal matrix D and a unitary matrix U with

$$T = UDU^\dagger = (U\sqrt{D}U^\dagger)(U\sqrt{D}U^\dagger)$$

That is, T has a square root which is itself non-negative and positive definite. The same result holds for separable Hilbert spaces.

Theorem 10. *Let T be non-negative and self-adjoint on a separable Hilbert space, then there exists a unique non-negative self adjoint operator \sqrt{T} such that $T = \sqrt{T}\sqrt{T}$.*

This follows from the spectral theorem or representation of C^* star algebras see [11] page 217. We note that given any operator $T \in \mathcal{L}(U, H)$ the operator TT^* is non-negative and self-adjoint. As a result we can make the following definition.

Definition 24. *Given a linear operator T between Hilbert spaces we define*

$$|T| := \sqrt{TT^*}.$$

5.6 Trace and Hilbert Schmidt operators

We consider in this subsection U, H to be separable Hilbert spaces, though much would remain valid for complex Hilbert spaces which are not necessarily separable with minimal adjustments.

Definition 25. *A linear operator $Q \in L(U)$ is said to be non-negative if*

$$\langle Qu, u \rangle \geq 0 \quad \forall u \in U.$$

Definition 26. *A linear operator $Q \in L(U)$ is said to be symmetric (or self-adjoint) if*

$$\langle Qu, v \rangle = \langle u, Qv \rangle \quad \forall u, v \in U.$$

Definition 27. A partial isometry is a linear map $Q : U \rightarrow H$ such that Q is an isometry on $\text{Ker}(Q)^\perp$.

Theorem 11 (Polar Theorem). Let $Q : U \rightarrow H$ be a linear operator between Hilbert spaces, then there is a unique partial isometry T such that

$$Q = T|Q|; \quad \ker(T) = \ker(Q).$$

Additionally $T^*Q = |Q|$.

Proof. Since $\|T\| = \||T|\|$ we may define $T(|Q|x) = Qx$ and $Ty = 0$ for $y \in \ker Q$. This can then be extended continuously to $U = \overline{\text{Im}(|Q|)} \oplus \text{Im}(|Q|)^\perp$. See [14] page 51. \square

The above theorem is the analogous of the fact that a complex number can be written as

$$z = e^{i \arg(z)} |z|.$$

Definition 28. A linear operator $Q \in L(U \rightarrow H)$ is said to have finite trace if for every orthonormal basis $\{e_k\}_{k=1}^\infty \subset U$

$$\|Q\|_1 := \sum_{k \in \mathbb{N}} \langle |Q|e_k, e_k \rangle < \infty.$$

It is also said that Q is trace class or nuclear. We write the space of such functions as $\mathcal{L}_1(U)$.

Despite the name, we cannot define the trace of a trace class operator Q unless $U = H$. In this case we say that the trace of Q is the quantity

$$\text{Tr}(Q) := \sum_{n=0}^{\infty} \langle Qe_n, e_n \rangle.$$

We note that $\text{Tr}(Q) \leq \|Q\|_1$ and that is is not enough for the trace of Q to exist for Q to be trace class as the sum may convergence but not absolutely. In the literature one may find many other equivalent definitions for trace class operators. We include the following list below which may be found in [11] page 92.

Proposition 30. Given two Hilbert spaces (not sure if separable would be necessary...) U, H and a linear operator $Q \in L(U, H)$ the following are equivalent.

1. Q is trace class.
2. For all orthonormal basis of U, H we have $\sum_{n=0}^{\infty} \langle Qe_n, v_n \rangle < \infty$.
3. Q is compact and $\sum_{n=0}^{\infty} \lambda_n < \infty$ where λ_n are the eigenvalues of $|Q|$.
4. Q is nuclear, that is there exists $\lambda_n \in \mathbb{R}, \ell_n \in H^*$ and $y_n \in H$ bounded sequences such that

$$T(x) = \sum_{n=0}^{\infty} \lambda_n \ell_n(x) y_n \quad \forall x \in H.$$

5. Q^* is nuclear.
6. $\sum_{n=0}^{\infty} \|Qe_n\| < \infty$ for some orthonormal basis.

Definition 29. We write $\mathcal{L}_1^+(U)$ for the space of trace class, positive semi-definite, symmetric operators.

Though in infinite dimension this definition could depend on the basis it will not do so for the class of linear operators we are interested in. This is because due to the following result all these operators are nuclear.

Proposition 31 (Hilbert Schmidt for \mathcal{L}_1^+). *If $Q \in \mathcal{L}_1^+(U)$ then there exists an orthonormal basis $e_k, k \in \mathbb{N}$, of U such that*

$$Qe_k = \lambda_k e_k, \quad \lambda_k \geq 0 \quad \forall k \in \mathbb{N}.$$

Proof. By definition of $\mathcal{L}_1^+(U)$ we have that Q is self-adjoint and compact (as it is nuclear). As a result we may apply the Hilbert Schmidt theorem 9 to obtain that

$$Qx = \sum_{k \in \mathbb{N}} \lambda_k \langle x, e_k \rangle e_k.$$

Finally, $\lambda_k \geq 0$ as Q is positive semi-definite. □

5.7 Elliptic SPDEs

In this section we define what is meant by a linear SPDE of the form $\mathcal{L}X - dW$. To do so we must first define white noise and thus generalized fields. The following definition is taken from [11] page 105

Definition 30. A generalized random field over a topological vector space V is a collection of random variables $\{\mathcal{X}(u)\}_{u \in V}$ such that the following hold

1. Linearity: $\mathcal{X}(u + av) = X(u) + aX(v)$ for all $u, v \in V$ and $a \in \mathbb{R}$.
2. Continuity: If $u_n \rightarrow u \in V$ then we have convergence in probability $\mathcal{X}(u_n) \xrightarrow{\mathbb{P}} \mathcal{X}(u)$.

A particular class of generalized fields are the Gaussian ones

Definition 31. We say that a 0 mean generalized Gaussian field \mathcal{B} over H is a collection of real valued Gaussian random variables $\{\mathcal{B}(f)\}_{f \in H}$ such that

1. $\mathbb{E}[\mathcal{B}(f)] = 0$ for all $f \in H$.
2. There exists $K \in \mathcal{L}_1^+(H)$ such that

References

- [1] E. Çinlar, E. Şınlar, Probability and stochastics, Vol. 261, Springer, 2011.
- [2] M. Taylor, Partial differential equations II: Qualitative studies of linear equations, Vol. 116, Springer Science & Business Media, 2013.
- [3] G. Lindgren, Stationary stochastic processes: theory and applications, CRC Press, 2012.
URL [https://www.amazon.co.uk/Stochastic-Processes-Applications-St
dp/1466557796](https://www.amazon.co.uk/Stochastic-Processes-Applications-Stationary/dp/1466557796)
- [4] S. Bochner, Monotone funktionen, stieltjessche integrale und harmonische analyse, Mathematische Annalen 108 (1) (1933) 378–410.
- [5] R. B. Ash, B. Robert, C. A. Doleans-Dade, A. Catherine, Probability and measure theory, Academic press, 2000.
URL [https://www.elsevier.com/books/
probability-and-measure-theory/ash/978-0-08-051487-1](https://www.elsevier.com/books/probability-and-measure-theory/ash/978-0-08-051487-1)

- [6] A. M. Yaglom, Correlation Theory of Stationary and Related Random Functions, Volume I: Basic Results, Vol. 131, Springer, 1987.
- [7] R. J. Adler, J. E. Taylor, et al., Random fields and geometry, Vol. 80, Springer, 2007.
URL https://www.dcs.warwick.ac.uk/~feng/papers/Random_Fields_and_Geometry_%5BAdler%5D.pdf
- [8] A. M. Yaglom, et al., Second-order homogeneous random fields, in: Proc. 4th Berkeley Sympos. Math. Statist. and Prob, Vol. 2, 1961, pp. 593–622.
- [9] B. Øksendal, Stochastic differential equations, in: Stochastic differential equations, Springer, 2003, pp. 65–84.
- [10] P. Whittle, Stochastic-processes in several dimensions, Bulletin of the International Statistical Institute 40 (2) (1963) 974–994.
- [11] S. V. Lototsky, B. L. Rozovsky, et al., Stochastic partial differential equations, Springer, 2017.
- [12] W. Liu, M. Röckner, Stochastic partial differential equations: an introduction, Springer, 2015.
- [13] M. Hairer, An introduction to stochastic pdes, arXiv preprint arXiv:0907.4178 (2009).
- [14] G. J. Murphy, C*-algebras and operator theory, Academic press, 2014.