

# Stuff studied

Liam Llamazares

February 2022

## 1 Gaussian Random variables

### 1.1 To Banach spaces

Let  $E$  be a Banach space and  $(\Omega, \mathcal{F}, \mathbb{P})$  a probability space

**Definition 1.** A Gaussian probability measure on  $E$  is a Borel measure  $\mu$  such that for all  $\ell \in E^*$  it holds that  $\ell^* \mu$  is a Gaussian measure on  $\mathbb{R}$ .

**Definition 2.** A Gaussian random variable to  $E$  is a measurable function  $W : (\Omega, \mathcal{F}) \rightarrow (E, \mathcal{B}(E))$  such that for all  $\ell \in E^*$  we have that  $\ell(W)$  is a Gaussian random variable.

**Proposition 1.** Let  $\mu, \nu$  be two finite measures on a separable Banach space  $E$ , then

$$\mu = \nu \iff \ell^* \mu = \ell^* \nu \quad \forall \ell \in E^*$$

*Proof.* See Hairer, use separability to show that the  $\sigma$ -algebra generated by cylinders is the same as the Borel  $\sigma$ -algebra. Since the two measure coincide on cylinders  $T^{-1}(\mathbb{R}^d)$  for  $T$  continuous (by knowledge that measures on  $\mathbb{R}^d$  are determined by their projections) also on the generated  $\sigma$ -algebra (by finiteness of measure).  $\square$

**Proposition 2.** Let  $\mu$  be a measure on a separable Banach space, then  $\mu$  is determined by it's Fourier transform

*Proof.* In the particular case  $\mathcal{B} = \mathbb{R}^n$ , if  $\varphi$  is a smooth function with compact support, it follows from Fubini's theorem and the invertibility of the Fourier transform that one has the identity

$$\int_{\mathbb{R}^n} \varphi(x) \mu(dx) = \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} \hat{\varphi}(\xi) e^{-i\xi \cdot x} d\xi \mu(dx) = \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} \hat{\varphi}(\xi) \hat{\mu}(-\xi) d\xi,$$

so that, since bounded continuous functions can be approximated by smooth functions,  $\mu$  is indeed determined by  $\hat{\mu}$ . The general case then follows immediately from the previous proposition.  $\square$

**Definition 3.** Given a Gaussian random measure  $\mu$  and a Gaussian random variable  $W$  we define the covariance of  $W$  to be  $C_\mu, C_W : E^* \times E^* \rightarrow \mathbb{R}$

$$C_\mu(\ell_1, \ell_2) := \text{Cov}_\mu[\ell_1, \ell_2]; \quad C_W(\ell_1, \ell_2) := C_{\mathbb{P}_W}(\ell_1, \ell_2) = \text{Cov}[\ell_1(W), \ell_2(W)]$$

Note that  $C_W$  is symmetric, positive definite and bi-linear. Furthermore we can define  $\tilde{C}_W : E^* \rightarrow E^{**}$  as  $\tilde{C}_W(\ell_1)(\ell_2)L = C_W(\ell_1, \ell_2)L$ .

**Proposition 3.** Let  $\mu$  be a Gaussian measure on a separable Banach space  $E$ , then  $C_\mu$  is continuous and  $\tilde{C}_\mu$  has image in  $E$

*Proof.* This follows from Fernique's theorem, see Hairer.  $\square$

**Definition 4.** The reproducing kernel Hilbert space associated with  $W$  is the completion of  $E^*$  with the norm

$$\|\ell\|_{R_W} := \sqrt{C_W(\ell, \ell)}$$

and is written  $R_W$

## 1.2 To Hilbert spaces

Let  $U$  be a separable Hilbert space, then we make the following definitions.

**Definition 5.** A measure  $\mu$  on  $(U, \mathcal{B}(U))$  is said to be Gaussian if for all  $u \in U$  there exist  $m(u), \sigma(u)$  such that, for  $\sigma(u) \neq 0$

$$\mu_{\langle \cdot, u \rangle}(A) = \int_A f(m(u), \sigma^2(u), s) ds \quad \forall A \in \mathcal{B}(\mathbb{R}).$$

Where  $f(m, \sigma^2, \cdot)$  is the density of a normal random variable with mean  $m$  and variance  $\sigma^2$ . And otherwise

$$\mu_{\langle \cdot, u \rangle} = \delta_{m(u)}.$$

**Definition 6.** The Fourier transform of  $\mu$  is the function  $\hat{\mu} : U \rightarrow \mathbb{R}$  defined by

$$\hat{\mu}(v) := \int_U e^{i\langle u, v \rangle} d\mu(v).$$

**Proposition 4.** *A measure  $\mu$  on  $(U, \mathcal{B}(U))$  is Gaussian if and only if*

$$\hat{\mu}(u) := \int_U e^{i\langle u, v \rangle_U} \mu(dv) = e^{i\langle m, u \rangle_U - \frac{1}{2}\langle Qu, u \rangle_U}, \quad u \in U,$$

where  $m \in U$  and  $Q \in L(U)$  is non-negative, symmetric, with finite trace. In this case  $\mu$  will be denoted by  $N(m, Q)$  where  $m$  is called the mean and  $Q$  is called the covariance (operator). The measure  $\mu$  is uniquely determined by  $m$  and  $Q$ . Furthermore, for all  $h, g \in U$

$$\begin{aligned} \int_U \langle x, h \rangle_U \mu(dx) &= \langle m, h \rangle_U, \\ \int_U (\langle x, h \rangle_U - \langle m, h \rangle_U) (\langle x, g \rangle_U - \langle m, g \rangle_U) \mu(dx) &= \langle Qh, g \rangle_U, \\ \int_U \|x - m\|_U^2 \mu(dx) &= \text{tr } Q. \end{aligned}$$

It follows from the above proposition that, the values of  $m(u), \sigma(u)$  in Definition 5 are

$$m(u) = \langle m, u \rangle; \quad \sigma^2(u) = \langle Qu, u \rangle.$$

Using this we can get the following representation theorem

**Proposition 5.** *(Representation of a Gaussian Random Variable) Let  $m \in U$  and  $Q \in L(U)$  be non-negative, symmetric, with  $\text{tr } Q < \infty$ . Let  $e_k$  be an orthonormal basis of  $U$  consisting of eigenvectors of  $Q$  with eigenvalues  $\lambda_k$ . Then a  $U$ -valued random variable  $X$  on a probability space  $(\Omega, \mathcal{F}, P)$  is Gaussian with  $P \circ X^{-1} = N(m, Q)$  if and only if*

$$X = \sum_{k \in \mathbb{N}} \sqrt{\lambda_k} \beta_k e_k + m \quad (\text{as objects in } L^2(\Omega, \mathcal{F}, P; U))$$

where  $\beta_k, k \in \mathbb{N}$ , are independent real-valued random variables with  $P \circ \beta_k^{-1} = N(0, 1)$  for all  $k \in \mathbb{N}$  with  $\lambda_k > 0$ . The series converges in  $L^2(\Omega, \mathcal{F}, P; U)$ .

The idea of the implication is to take  $\beta_k$  to be

$$\beta_k := \begin{cases} = \frac{\langle X, e_k \rangle - \langle m, e_k \rangle}{\sqrt{\lambda_k}} & \text{if } \lambda_k > 0 \\ \equiv 0 \in \mathbb{R} & \text{else.} \end{cases}$$

Then we get that  $\beta_k \sim \mathcal{N}(0, 1)$  by the previous. It remains to show that they are independent.

The idea of the reciprocal is as follows. By construction we have that the finite sums are Gaussian. Furthermore, the infinite sum converges in  $L^2(\Omega, \mathcal{F}, P; U)$  so it is also Gaussian (use the characteristic function and dominated convergence theorem and independence and the below lemma). The calculation of the mean and the covariance is a calculation using (4).

**Lemma 1.** *Let  $Q$  be non-negative, symmetric and with finite trace and consider an orthonormal family  $\{e_k, \lambda_k\}$ . Then if  $\beta_k \sim \mathcal{N}(0, 1)$  we have*

$$X_k := \sqrt{\lambda_k} \beta_k e_k + m_k \sim \mathcal{N}(m_k, P_k Q)$$

Where  $P_k$  is the projection onto  $e_k$ .

*Proof.* By considering the definition of Gaussian random variable we obtain that it is Gaussian with  $m(u) = \langle m_k, u \rangle$  and

$$\sigma^2(u) = \lambda_k \langle e_k, u \rangle^2 = \langle Q e_k, u \rangle \langle e_k, u \rangle = \langle P_k Q u, u \rangle.$$

Since we have seen that

$$m(u) = \langle m, u \rangle; \quad \sigma^2(u) = \langle Q u, u \rangle.$$

This concludes the proof. □

As a corollary of the representation theorem we get the following.

**Corollary 1.** *Given  $m, Q$  there exists a Gaussian random variable  $X \sim \mathcal{N}(m, Q)$  and a Gaussian random measure  $\mu = \mathcal{N}(m, Q)$ .*

By using the definition of Gaussian measure it is possible to define the meaning of a Wiener process taking values in  $U$ .

**Definition 7.** *A  $U$  valued standard  $Q$ -Wiener process is a stochastic process*

$$W : (\Omega, \mathcal{F}, \mathbb{P}) \rightarrow U$$

*such that*

1.  $W(0) = 0$ .
2.  $W$  has  $\mathbb{P}$  almost everywhere continuous trajectories.

3. The increments  $W(t_1), W(t_2) - W(t_1), W(t_3) - W(t_2), \dots$  are independent random variables
4.  $W(t) - W(s) \sim \mathcal{N}(0, (t-s)Q)$  for all  $t \geq s$ . That is, for all  $A \in \mathcal{B}(\mathbb{R})$  and all  $u \in U$

$$\mathbb{P}(\langle W(t) - W(s), u \rangle \in A) = \frac{1}{\sqrt{2\pi \langle Qu, u \rangle}} \int_A \exp\left(-\frac{x^2}{2 \langle Qu, u \rangle}\right) dx.$$

**Lemma 2** (Doob's maximal Martingale inequality). *Let  $\{X_k\}_{k=1}^\infty$  be a sub-martingale. Then it holds that*

$$\left\| \max_{k \in \{1, \dots, n\}} X_k \right\|_{L^p(\Omega)} \leq \frac{p}{p-1} \|X_n\|_{L^p(\Omega)}$$

As a consequence, if  $X_t, t \in [0, T]$  is a left (or right) continuous semi-martingale then

$$\left\| \max_{t \in [0, T]} X_t \right\|_{L^p(\Omega)} \leq \frac{p}{p-1} \|X_T\|_{L^p(\Omega)}.$$

*Proof.* We only show how to get from the discrete to the continuous case. Take

$$F_j := \{0, t_0, \dots, t_j, T\}; \quad \{t_j\}_{j=1}^\infty = \mathbb{Q} \cap [0, T]$$

Then  $F_j \subset F_{j+1}$  so  $\max_{t \in F_j} X_t$  is an increasing sequence. And by continuity it converges to  $\max_{t \in [0, T]} X_t$ . So by the discrete case.

$$\left\| \max_{t \in [0, T]} X_t \right\|_{L^p(\Omega)} = \lim_{j \rightarrow \infty} \left\| \max_{t \in F_j} X_t \right\|_{L^p(\Omega)} \leq \lim_{j \rightarrow \infty} \frac{p}{p-1} \|X_T\|_{L^p(\Omega)} = \frac{p}{p-1} \|X_T\|_{L^p(\Omega)}$$

□

As a result of this and the previous proposition we get the following representation theorem

**Proposition 6.** (*Representation of a  $Q$ -Wiener Process*) *Let  $Q \in L(U)$  be non-negative, symmetric, with  $\text{tr } Q < \infty$ . Let  $e_k$  be an orthonormal basis of  $U$  consisting of eigenvectors of  $Q$  with eigenvalues  $\lambda_k$ . Then a  $U$ -valued process  $W$  on a probability space  $(\Omega, \mathcal{F}, P)$  is a  $Q$ -Wiener Process if and only if*

$$X(t) = \sum_{k \in \mathbb{N}} \sqrt{\lambda_k} \beta_k(t) e_k \quad (\text{as objects in } L^2(\Omega, \mathcal{F}, P; U))$$

where  $\beta_k, k \in \mathbb{N}$ , are independent real-valued Wiener processes on  $(\Omega, \mathcal{F}, P)$  for all  $k \in \mathbb{N}$  with  $\lambda_k > 0$ . Furthermore, the series converges in  $L^2(\Omega \rightarrow C([0, T] \rightarrow U))$ .

*Proof.* To prove the implication we note that, by previous results, it holds that for

$$\beta_k := \begin{cases} = \frac{\langle W, e_k \rangle}{\sqrt{\lambda_k}} & \text{if } \lambda_k > 0 \\ \equiv 0 \in \mathbb{R} & \text{else.} \end{cases}$$

we have  $\beta_k(t) \sim \mathcal{N}(0, t)$ . For the same reason

$$\beta_k(t) - \beta_k(s) = \frac{\langle W(t) - W(s), e_k \rangle}{\sqrt{\lambda_k}}$$

follows a  $\mathcal{N}(0, t - s)$ . The continuity of  $\beta_k$  is also clear by the continuity of  $W_k$  and they start at 0. It remains to check independence this is technical.

To prove the reverse implication we note that convergence holds in the above space by Doob's maximal Martingale inequality

$$\begin{aligned} E \left( \sup_{t \in [0, T]} \|W^N(t) - W^M(t)\|_U^2 \right) &= E \left( \sup_{t \in [0, T]} \sum_{k=M+1}^N \lambda_k \beta_k^2(t) \right) \\ &\leq \sum_{k=M+1}^N \lambda_k E \left( \sup_{t \in [0, T]} \beta_k^2(t) \right) \leq 4 \sum_{k=M+1}^N \lambda_k \sup_{t \in [0, T]} E(\beta_k(t)^2) \leq 4T \sum_{k=M+1}^N \lambda_k. \end{aligned}$$

The independence of the increments and that  $W(t) - W(s) \sim \mathcal{N}(0, Q(t - s))$  is a calculation. The fact that the process starts at 0 is also clear as  $\beta_k(0) = 0$ .  $\square$

We now define the filtrations with respect to which we will consider adapted Wiener processes. These verify two properties, completeness and continuity.

**Definition 8.** A normal filtration  $\mathcal{F}_t$  is a filtration such that

1. All sets with measure 0 are in  $\mathcal{F}_0$ .
2.  $\mathcal{F}_s = \mathcal{F}_{s+} = \bigcap_{s < t} \mathcal{F}_t$  for all  $s \in [0, T]$ .

The purpose of the first requirement is so that adaptations of adapted processes are also adapted. Though it is not necessary to construct the stochastic integral.

**Definition 9.** A  $Q$ -Wiener process  $W(t)$  is called a  $Q$ -Wiener process with respect to a normal filtration  $\mathcal{F}_t$  if

1.  $W(t)$  is adapted to  $\mathcal{F}_t$ .
2.  $W(t) - W(s)$  is independent of  $\mathcal{F}_s$  for all  $0 \leq s \leq t \leq T$ .

Given any  $Q$ -Wiener process  $W$  it can be shown that it is a Wiener process wrt the filtration

$$\mathcal{F}_s := \bigcap_{s < t} \sigma(\tilde{\mathcal{F}}_t \cup \mathcal{N}_0); \quad \tilde{\mathcal{F}}_t = \sigma(\{W_r : r \leq t\}); \quad \mathcal{N}_0 = \{A \in \mathcal{F} : \mathbb{P}(A) = 0\}$$

## 2 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 10.** 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 11.** 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  $L^p(\Omega \rightarrow X)$ . □

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 3.** *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  $\omega$  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  $\mathbb{R}$

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

$\square$

As a corollary we obtain the following

**Corollary 2.**  *$\mathcal{A}$  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 12.** 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 3.** 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

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

### 3 Martingales

One can use the Bochner integral to extend the existence of conditional probabilities to random variables in  $L^1(\Omega \rightarrow X)$ . This is done by proving existence for simple functions and then taking limits as every Bochner integrable function is limit of simple functions. This allows us to define the concept of Martingale.

**Definition 13.** Let  $M(t), t \geq 0$ , be a stochastic process on  $(\Omega, \mathcal{F}, \mathbb{P})$  with values in a separable Banach space  $E$ , and let  $\mathcal{F}_t, t \geq 0$ , be a filtration on  $(\Omega, \mathcal{F}, P)$ . The process  $M$  is called an  $\mathcal{F}_t$ -martingale, if:

1.  $E(\|M(t)\|) < \infty$  for all  $t \geq 0$ ,
2.  $M(t)$  is  $\mathcal{F}_t$ -measurable for all  $t \geq 0$ ,
3.  $E(M(t)|\mathcal{F}_s) = M(s), \mathbb{P}$ -a.s. for all  $0 \leq s \leq t < \infty$ .

Using that given a separable Hilbert space  $E$  there exists a sequence of linear operators  $l_n$  such that

$$\|z\| = \sup_n l_n(z) \quad \forall z \in E.$$

We can deduce using the linearity of the Bochner integral that, if  $M$  is a martingale then  $\|M\|$  is a submartingale. Furthermore, by the conditional version of Jensen's inequality it holds that if  $\varphi$  is a convex lower semi continuous function or convex and increasing and such that  $\varphi(M(t)) \in L^1(\Omega \rightarrow E)$  for all  $t$  then  $\varphi(M(t))$  is a sub-martingale. Now applying Doob's maximal inequality gives that

**Theorem 1** (Maximal Inequality). *Let  $p > 1$  and let  $E$  be a separable Banach space. If  $M(t), t \in [0, T]$ , is a right-continuous  $E$ -valued  $\mathcal{F}_t$ -martingale, then*

$$\begin{aligned} \left( E \left( \sup_{t \in [0, T]} \|M(t)\|^p \right) \right)^{\frac{1}{p}} &\leq \frac{p}{p-1} \sup_{t \in [0, T]} (E(\|M(t)\|^p))^{\frac{1}{p}} \\ &= \frac{p}{p-1} (E(\|M(T)\|^p))^{\frac{1}{p}} \end{aligned}$$

*Proof.* This follows by using that  $\|M(t)\|$  is a sub-martingale and Doob's maximal inequality.  $\square$

**Corollary 4.** *Let  $M$  be a continuous martingale to a separable Banach space  $E$ . Then the following are equivalent*

- $M \in \hat{L}^\infty([0, T] \rightarrow \hat{L}^2(\Omega \rightarrow E))$
- $M \in \hat{L}^2(\Omega \rightarrow \hat{L}^\infty([0, T] \rightarrow E))$
- $\mathbb{E}[\|M(T)\|^2] < \infty$

A useful space of Martingales is as follows

**Definition 14.** *Let  $E$  be a separable Banach space, then we define*

$$\mathcal{M}_T^2(E) := \{ \text{continuous martingales } M : \mathbb{E}[\|M(T)\|^2] < \infty \}$$

*and give it the norm*

$$\|M\|_{\mathcal{M}_T^2(E)} := \mathbb{E}[\|M(T)\|^2].$$

By Theorem (1) we have that

$$\mathcal{M}_T^2(E) \subset \hat{L}^\infty([0, T] \rightarrow \hat{L}^2(\Omega \rightarrow E)) \cap \hat{L}^2(\Omega \rightarrow \hat{L}^\infty([0, T] \rightarrow E)).$$

and that any of the norms of these spaces is equivalent to the one set on  $\mathcal{M}_T^2(E)$ . This is useful in the following result

**Proposition 9.** *Let  $E$  be a separable Banach space, then  $\mathcal{M}_T^2(E)$  is a Banach space.*

*Proof.* By the previous observation and the Fischer-Riesz theorem,  $\mathcal{M}_T^2(E)$  is a subspace of a Hilbert space. As a result it is sufficient to show that it is closed. Let  $M_n$  converge to  $M$ . Then, by the equivalence of the norms we have that  $M_n(t) \rightarrow M(t) \in \hat{L}^1(\Omega \rightarrow E) \subset \hat{L}^2(\Omega \rightarrow E)$  so that for all  $A \in \mathcal{F}_s$

$$\int_A M(s) d\mathbb{P} = \lim_{n \rightarrow \infty} \int_A M_n(s) d\mathbb{P} = \lim_{n \rightarrow \infty} \int_A M_n(t) d\mathbb{P} = \int_A M(t) d\mathbb{P}.$$

This shows that  $M$  is a martingale. Furthermore, as was seen in the proof of the Fischer Riesz theorem, there exists a subsequence  $M_{n_k}$  such that

$$\lim_{n \rightarrow \infty} M_{n_k}(\omega) = M(\omega) \in \hat{L}^\infty([0, T] \rightarrow E) \quad a.e. \quad \omega \in \Omega$$

Since  $M_{n_k}(\omega)$  are continuous and continuity is preserved by uniform limits this proves that  $M$  is continuous almost everywhere. This concludes the proof.  $\square$

**Proposition 10.** *Let  $W(t)$  be a  $E$  valued  $\Sigma$ -Wiener process with respect to a filtration  $\mathcal{F}_t$ . Then  $W(t) \in \mathcal{M}_T^2(E)$ .*

*Proof.* It is a martingale as it is adapted and, given  $A \in \mathcal{F}_s$  and  $u \in E$ , by linearity of the integral and Independence of  $W(t) - W(s)$  with  $\mathcal{F}_s$

$$\begin{aligned} \left\langle \int_A W(t) - W(s) d\mathbb{P}, u \right\rangle &= \int_A \langle W(t) - W(s), u \rangle d\mathbb{P} \\ &= \mathbb{P}(A) \mathbb{E}[\langle W(t) - W(s), u \rangle] = 0 \end{aligned}$$

As a result

$$\int_A W(t) d\mathbb{P} = \int_A W(s) d\mathbb{P} = 0 \quad \forall A \in \mathcal{F}_s \implies \mathbb{E}_s[W(t)] = W(s).$$

Finally, we have that  $\mathbb{E}[W(t)^2] = t < \infty$  for all  $t$  and  $W$  is continuous by construction. This concludes the proof.  $\square$

## 4 Operators

**Proposition 11.** *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 12.** *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 13** (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 14** (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.  $\square$

### 4.1 Compact operators

Let  $X, Y$  be normed spaces.

**Definition 15.** *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 15.** *Let  $T : X \rightarrow Y$  have finite range, then  $T$  is compact  $\iff$  it is continuous.*

**Proposition 16.** *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 17.**  $T$  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 18.** 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 19.** 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 20.** 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.  $\square$

## 4.2 Fredholm alternative

**Proposition 21.** *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 is the identity here so this space must have finite dimension  $\square$

## 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 22.** *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  $S$ .  $\square$

## 4.4 Specter of a linear operator

**Definition 16.** *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 17.** Let  $T \in \mathcal{L}(E)$  with  $E$  a Banach space, then we say that the eigenvalues of  $T$  are the  $\lambda$  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 23.** 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  $\mu$ . 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 18.** The spectral radius of  $T$  is defined as

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

**Proposition 24.** 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 25.** 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  $\lambda$  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  $\lambda$  is in the spectrum □

## 5 Self-adjoint operators

**Proposition 26.** *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 27.** *Let  $T$  be a self-adjoint operator on a Hilbert space, then its spectrum is non-empty (and contains  $\|T\|$  or  $-\|T\|$ )*

**Proposition 28.** *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  $T|_{E^\perp}$  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 2** (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  $\sigma_P$  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 3.** *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 [?] 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 19.** *Given a linear operator  $T$  between Hilbert spaces we define*

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

## 6 Trace and Hilbert Schmidt operators

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

**Definition 20.** *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 21.** *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 22.** A partial isometry is a linear map  $Q : U \rightarrow H$  such that  $Q$  is an isometry on  $\text{Ker}(Q)^\perp$ .

**Theorem 4** (Polar Theorem). Let  $Q : U \rightarrow H$  be a linear operator between Hilbert spaces, then there is a unique partial isometry  $T$  such that

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

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

*Proof.* Since  $\|T\| = \||T|\|$  we may define  $T(|Q|x) = Qx$  and  $Ty = 0$  for  $y \in \text{ker } Q$ . This can then be extended continuously to  $U = \overline{\text{Im}(|Q|)} \oplus \text{Im}(|Q|)^\perp$ . See [?] 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 23.** 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 [?] page 92.

**Proposition 29.** 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.  $\sqrt{|Q|}$  is a Hilbert Schmidt operator.
3. For all orthonormal basis of  $U, H$  we have  $\sum_{n=0}^{\infty} \langle Qe_n, v_n \rangle < \infty$ .
4.  $Q$  is compact and  $\sum_{n=0}^{\infty} \lambda_n < \infty$  where  $\lambda_n$  are the eigenvalues of  $|Q|$ .
5.  $Q$  is nuclear, that is there exists  $\lambda_n \in \mathbb{R}, \ell_n \in H^*$  and  $y_n \in H$  bounded sequences such that

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

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

**Definition 24.** 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 30** (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 2 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. □

We fix  $U, H$  to be separable Hilbert spaces

**Definition 25.** Let  $e_k$  be an orthonormal basis of  $U$ . Then  $A \in L(U, H)$  is said to be a Hilbert-Schmidt operator if

$$\|A\|_{L_2(U, H)} := \sum_{k \in \mathbb{N}} \langle Ae_k, Ae_k \rangle < \infty.$$

**Proposition 31.** The space of Hilbert-Schmidt operators  $L_2(U, H)$  is a separable Hilbert space of compact operators with the inner product

$$\langle A, B \rangle := \sum_{k \in \mathbb{N}} \langle Ae_k, Be_k \rangle.$$

Furthermore the inner product is independent of the choice of basis  $e_k$  of  $U$ .

*Proof.* The completeness follows from that of the analogously defined  $L_1(U, H)$  [?, Corollary 16.25, p.154] and Fatou's lemma.

By Parseval's identity given a basis  $f_k$  of  $H$  it holds that

$$\sum_k \|Ae_k\|^2 = \sum_k \sum_j |\langle Ae_k, f_j \rangle|^2 = \sum_j \sum_k |\langle A^* f_j, e_k \rangle|^2 = \sum_j \|A^* f_j\|^2$$

This proves the last part of the proposition as  $\|T\| = \|T^*\|$ . To prove that it is separable we show that

$$f_j \otimes e_k := \langle e_k, \cdot \rangle f_j$$

is an orthonormal basis (the orthonormality is a calculation) and that is is a basis follows from the fact that  $\langle A, f_j \otimes e_k \rangle = \langle f_j, Ae_k \rangle$ . So if all of these terms are 0 then  $A = 0$  as

$$Ax = \sum_{k,j} \langle x, e_k \rangle \langle Ae_k, f_j \rangle f_j = \sum_{k,j} \langle x, e_k \rangle \langle A, f_j \otimes e_k \rangle f_j.$$

□

**Proposition 32.** If  $A \in L^2(U, H)$  then it is also compact

*Proof.* This follows from the fact that by taking  $A_n := A \circ \Pi_n$  where  $\Pi_n$  is the projection onto the finite dimensional space generated by  $\{e_1, \dots, e_n\}$

$$\|Ax - A_n x\| = \left\| \sum_{n=0}^{\infty} \langle x, e_n \rangle Ae_n \right\| \leq \|x\|^2 \|A\|_{L_2(U, H)}$$

and the completeness of the space of compact operators to Banach spaces. □

**Proposition 33.** *Given  $A \in L_2(U, H)$  and  $T_1 \in L(H_1, H), T_2 \in L(H, H_2)$  with  $H_1, H_2$  separable Hilbert spaces. Then we have that*

- $\|A\|_{L_2(U, H)} = \|A^*\|_{L_2(H, U)}$  where  $A^*$  is the adjoint of  $A$ .
- $A \circ T_1 \in L_2(U, H_2)$  and  $T_2 \circ A \in L_2(U, H_2)$ .

**Proposition 34.** *Let  $\Sigma \in L(U, U)$  be nonnegative and symmetric. Then there exist a unique element  $\Sigma^{\frac{1}{2}}$  verifying that*

$$\Sigma^{\frac{1}{2}} \circ \Sigma^{\frac{1}{2}} = \Sigma.$$

*Furthermore, if  $\Sigma$  has finite trace then  $\Sigma^{\frac{1}{2}} \in L^2(U, U)$  and  $\left\| \Sigma^{\frac{1}{2}} \right\|_{L_2(U, U)}^2 = \text{tr}(\Sigma)$ .*

The proof can be found in [?, p.196].

## 7 Reproducing kernel

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

**Definition 26.** *A reproducing kernel space is a Hilbert space  $H \subset \mathbb{C}^S$  where  $S$  is some set and such that point evaluation*

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

*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 itself with kernel the “identity”. To give another example, by Sobolev embedding theory,  $H^s(\mathbb{R}^d)$  are reproducing kernel spaces for  $s > \frac{d}{2}$ .

**Definition 27.** A real valued function  $K : S \times S$  is said to be a positive-definite Kernel if  $\{K(t_i, t_j)\}_{i,j=1}^n$  is a Hermitian, non-negative 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 real valued process is a kernel.

**Theorem 5** (Existence and uniqueness of reproducing kernels and their spaces). Given a reproducing kernel Hilbert space  $H$  there exists a unique 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

$$K_H(s, t) = \langle K_H(s, \cdot), K_H(t, \cdot) \rangle. \quad (1)$$

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  $H_K$ .  $\square$

The equality (1) 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 operator (kernel)* associated with  $\mu$ . By setting  $S = V'$  we may apply the previous to deduce that there exists a Hilbert space  $H_\mu$  whose reproducing kernel is  $C_\mu$ . We call this space the *reproducing kernel Hilbert space of  $\mu$* . If  $\mu$  is the distribution of some  $V$  valued random variable  $X$  then we also say that  $H_\mu$  is the *reproducing kernel Hilbert space of  $X$* .

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

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

**Theorem 6.** *Let  $H$  be a separable Hilbert space and identify  $H$  with  $H'$ , then:*

1. *Given a Gaussian measure  $\mu$  on a Hilbert space  $(H, \mathcal{B}(H))$ , 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  $\mu$  on  $(H, \mathcal{B}(H))$  such that  $C_\mu(u, v) = \langle Ku, v \rangle$ .*

*Proof.* • 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.

- 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  $\mu$  is a Gaussian measure and, by the transfer theorem and 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 35.** *Let  $E$  be a separable Banach space, then  $\mu$  is a Gaussian measure on  $E$  iff for every linear function  $\ell : E \rightarrow \mathbb{R}$*

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

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

*Proof.* By theory of 1-dimensional Gaussian random variables we have that  $\ell^* \mu$  is a Gaussian measure if and only if

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

Where, by the linearity of the Bochner integral,

$$\begin{aligned} a &= \int_E x \ell^* \mu(dx) = \int_E \ell(x) \mu(dx) = \ell(m) \\ \sigma &= \int_E x^2 \ell^* \mu(dx) = \int_E \ell(x) \ell(x) \mu(dx) = C_\mu(\ell, \ell). \end{aligned}$$

□

Furthermore, the Fourier transform uniquely characterizes the measure, as a result we obtain that a Gaussian measure on a separable Banach space  $E$  is uniquely characterized by its mean and covariance operator. However, note that there may in general be infinite Gaussian measures on  $E$  with the same mean and covariance operator. In fact, in the previous theorem we constructed one for each choice of orthonormal basis.

## 8 Construction of the stochastic integral

We fix  $U, H$  to be separable Hilbert spaces and  $E$  a separable Banach space. The way to construct the Stochastic integral is as follows

1. Construct for elementary processes  $Int : \mathcal{E} \rightarrow \mathcal{M}_T^2(H)$ .
2. It is a isometry so it extends uniquely and isometrically to  $\bar{\mathcal{E}}$ .
3. Characterise  $\bar{\mathcal{E}}$ .

In the case where  $U = \mathbb{R}^n, H = \mathbb{R}^m$  we have that elementary processes would be defined piece wise in time as some matrices in  $\mathbb{R}^{m \times n}$ . In this case we have similarly that

**Definition 28** (Elementary Process). *An  $L(U, H)$ -valued process  $\Phi(t)$ , on  $(\Omega, \mathcal{F}, P)$  with respect to  $\mathcal{F}_t$ , is said to be elementary if there exist  $0 = t_0 < \dots < t_k = T, k \in \mathbb{N}$ , such that*

$$\Phi(t) = \sum_{m=0}^{k-1} \Phi_m 1_{(t_m, t_{m+1}]}(t), \quad t \in [0, T]$$

Where  $\Phi_m : \Omega \rightarrow L(U, H)$  is measurable between  $\mathcal{F}_{t_m}$  and the Borel  $\sigma$ -algebra on  $L(U, H)$  and takes only a finite amount of values. We write the space of these processes by  $\mathcal{E} = \mathcal{E}(\Omega, \mathcal{F}, U, H)$ .

**Definition 29** (Stochastic integral for elementary processes). *Given  $\Phi \in \mathcal{E}$  and a  $U$  valued  $\Sigma$  Wiener process we define its stochastic integral as*

$$\text{Int}(\Phi)(t) \equiv \int_0^t \Phi(s) dW(s) := \sum_{m=0}^{k-1} \Phi_m(W(t \wedge t_{m+1}) - W(t_m \wedge t))$$

We note that if  $t < t_m$  then  $t \wedge t_{m+1} = t_m \wedge t$  so the above definition corresponds to transforming the increments of  $W$  up to time  $t$  using  $\Phi$ .

**Proposition 36.** *Given  $\Phi \in \mathcal{E}$  and a  $U$  valued  $\Sigma$  Wiener process*

$$\int_0^t \Phi(s) dW(s) \in \mathcal{M}_T^2(H)$$

*Proof.* The continuity in  $t$  is clear as  $W(t)$  is continuous and so are  $\Phi_m$ . Square integrability follows from this property holding for  $W(t)$  and the fact that, since  $\Phi_m$  takes only a finite amount of values in  $L(U, H)$ , the norm  $\Phi_m(\omega)$  is bounded. The fact that it is a Martingale is a technical calculation using the linearity of the integral and that  $\Phi_m$  takes only a finite amount of values in  $L(U, H)$ .  $\square$

**Proposition 37** (Ito's isometry for elementary processes). *Given  $\Phi \in \mathcal{E}$  it holds that*

$$\left\| \int_0^\cdot \Phi(s) dW(s) \right\|_{\mathcal{M}_T^2(H)}^2 = E \left( \int_0^T \left\| \Phi(s) \circ \Sigma^{\frac{1}{2}} \right\|_{L_2(U, H)}^2 ds \right) =: \|\Phi\|_{\mathcal{E}}^2$$

The proof is a very technical calculation. We then identify to elementary processes if the seminorm  $\|\cdot\|_T$  is equal. We note that this does not necessarily implt that the processes are equal  $dt_{[0, T]} \otimes \mathbb{P}$  almost everywhere. But rather that they correspond on  $\Sigma^{\frac{1}{2}}(U)$ ,  $dt_{[0, T]} \otimes \mathbb{P}$  almost everywhere.

**Corollary 6.** *The transformation*

$$\text{int} : \mathcal{E} \rightarrow \mathcal{M}_T^2(E).$$

*is an isometric linear transformation*

We now characterize  $\bar{\mathcal{E}}$ . We need the following definitions

**Definition 30.** Let  $H'$  be a separable Hilbert space and define

$$\begin{aligned}\mathcal{P}_T &:= \sigma\left(\{(s, t] \times F_s : 0 \leq s < t \leq T, F_s \in \mathcal{F}_s\} \cup \{\{0\} \times F_0 : F_0 \in \mathcal{F}_0\}\right) \\ &= \sigma(Y : \Omega_T \rightarrow \mathbb{R} : Y \text{ is left-continuous and adapted to } \mathcal{F}_t, t \in [0, T]).\end{aligned}$$

Then  $Y : \Omega \rightarrow H'$  is said to be  $H'$  predictable if it is  $\mathcal{P}_T \rightarrow \mathcal{B}(H')$  measurable.

**Definition 31.** Let  $H$  be a separable Hilbert space and  $\Sigma$  non-negative, symmetric, with finite trace. Then we define the separable Hilbert space

$$L_2^0 := L_2(U_0, H); \quad (U_0, \langle \cdot, \cdot \rangle_0) := \left( \Sigma^{\frac{1}{2}}(U), \left\langle \Sigma^{-\frac{1}{2}} \cdot, \Sigma^{-\frac{1}{2}} \cdot \right\rangle_U \right).$$

Which is the space of Hilbert Schmidt operators between  $U_0$  and  $H$ .

It can be shown that

$$\|f\|_{L_2^0} = \left\| f \circ \Sigma^{\frac{1}{2}} \right\|_{L_2(U \rightarrow H)}.$$

As a result, if we define

$$\|\Phi\|_T = \left( E \left( \int_0^T \|\Phi(s)\|_{L_2^0}^2 ds \right) \right)^{\frac{1}{2}}.$$

Then, by the definition in 37,  $\|\Phi\|_T = \|\Phi\|_{\mathcal{E}}$  for all  $\Phi \in \mathcal{E}$ . Furthermore, we can prove the following.

**Theorem 7.** The completion  $\bar{\mathcal{E}}$  of the elementary processes is given by

$$\begin{aligned}\mathcal{N}_W^2(0, T; H) &:= \{ \Phi : [0, T] \times \Omega \rightarrow L_2^0 : \Phi \text{ is predictable and } \|\Phi\|_T < \infty \} \\ &= L^2([0, T] \times \Omega, \mathcal{P}_T, dt \otimes P; L_2^0).\end{aligned}$$

Let  $E$  be a separable Banach space, then  $\mu$  is a Gaussian measure on  $E$  iff for every linear function  $\ell : E \rightarrow \mathbb{R}$

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

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

## References