Stochastic Partial Differential Equations

Based on Lectures by Dr. A. Chandra and Dr. G. Cannizzaro

November 28, 2019

Contents

1. Gaussian Measure Theory **3**
1.1. Introduction . . . 3
1.2. Analysis of Banach-valued random variables . . . 3
1.2.1. Two approaches . . . 3
1.2.2. Marginals characterise measures . . . 5
1.2.3. Mean and covariance . . . 6
1.2.4. Fourier transforms . . . 7
1.3. Gaussian measures . . . 7
1.3.1. Basic definition and mean-covariance characterisation . . . 7
1.3.2. Integrability of Gaussian measures . . . 8
1.4. Existence of Gaussian Measures . . . 9
1.4.1. Cameron-Martin theory . . . 12

A. Miscellaneous Theorems **13**

1. Gaussian Measure Theory

1.1. Introduction

The driving motivation throughout this chapter is defining a Banach space valued Wiener process and defining how to integrate against it. Recall a $\mathbb{R}^n$-valued stochastic process $(W_t)_{t\geq 0}$ is a Wiener process if and only if it is a continuous Gaussian process with mean 0 and covariance function $C(s,t) = (s \wedge t)I$ where I is the $d \times d$ identity matrix. This is a nice short definition, and it's easy to take for granted the amount of measure theory it took to get to this definition. Let's spell some of those steps out.

1. We developed tools like mean, covariance and Fourier transforms to analysis real random variables.
2. We defined what a Gaussian random variable was and used our tools to analyse their properties.
3. We showed existence of a Gaussian process with the correct covariance structure.
4. We proved a continuity criterion showing our Gaussian process has a continuous modification.

For a Banach space $\mathcal{B}$-valued stochastic process, we currently have none of these tools. The goal of this chapter is to remedy that.

1.2. Analysis of Banach-valued random variables

1.2.1. Two approaches

Throughout these notes $\mathcal{B}$ be a real Banach space. Usually $\mathcal{B}$ will also be seperable, but we'll try and say explicitly when this is the case. Let μ be a Borel probability measure on $\mathcal{B}$ and X be a $\mathcal{B}$-valued random variable such that $X \sim \mu$. We'd like recover tools like means and Fourier transforms. Let's start with the mean, say m. From the finite dimensional case it's tempting to define

$$m := \int_{\mathcal{B}} x \, \mu(\mathrm{d}x). \tag{1.1}$$

However this integral cannot be understand as a Lebesgue integral. After all, x doesn't take values in $\mathbb{R}^n$. At this point we have two options.

Option one: the Bochner integral

We could try and rigorously define this integral. The method of integration used for $\mathcal{B}$-valued functions are Bochner integrals. The construction of the Bochner integral follows the same method as the Lebesgue integral; we define the integral for a class of simple functions and extend the integral by approximation. Rather than going into details on the construction, we'll list off some of the properties of the Bochner integral. Firstly let $(E, \mathcal{E}, \nu)$ be a probability space and $f : E \to \mathcal{B}$ be measurable and $\mathcal{B}$ be seperable[1]. Then f is integrable if and only if

$$\int_E \|f(x)\|_E \, \nu(\mathrm{d}x) < \infty. \tag{1.2}$$

The Bochner integral behaves nicely with bounded linear operators. Let $\hat{\mathcal{B}}$ be another Banach space and $L : \mathcal{B} \to \hat{\mathcal{B}}$ be a bounded linear operator. Then the Bochner integral can be exchanged with L, i.e. for all integrable $f : E \to \mathcal{B}$,

$$\int_E L(f(x)) \, \nu(\mathrm{d}x) = L\left(\int_E f(x) \, \nu(\mathrm{d}x)\right). \tag{1.3}$$

Hence if we assume $\int_E \|x\| \, \mu(\mathrm{d}x) < \infty$ then the mean

$$\bar{m}_\mu := \int_E x \, \mu(\mathrm{d}x) \tag{1.4}$$

can be defined using the Bochner integral. This approach is valuable in how it carrys the intuition from the finite dimensional case. In addition, it makes clear that the mean is an element of $\mathcal{B}$ without the use of reflexivity (as we will see in the next approach).

Option two: one-dimensional marginals

Alternatively, we've already developed a rich theory of Lebesgue integration for $\mathbb{R}$-valued functions. What if we instead used this theory? This would require a 'natural' way to map $\mathcal{B}$ to $\mathbb{R}$. Since $\mathcal{B}$ is a normed vector space, we have exactly that: the space of bounded linear functionals $\mathcal{B}^*$. For any $l \in \mathcal{B}^*$, $\mathbb{E}[l(X)]$ is just given by a Lebesgue integral. So to define the mean, just gather all these values for diffent $l \in \mathcal{B}^*$. Following this idea, we define the mean as $m : \mathcal{B}^* \to \mathbb{R}$ where

$$m(l) := \mathbb{E}[l(X)] = \int_{\mathcal{B}} l(x) \, \mu(\mathrm{d}x). \tag{1.5}$$

How can we recover a mean $\bar{m}_\mu \in \mathcal{B}$? Suppose that $\int_E \|x\| \, \mu(\mathrm{d}x) < \infty$. We can then prove that $m_\mu : \mathcal{B}^* \to \mathbb{R}$ is a bounded linear functional, in other words $m_\mu \in \mathcal{B}^{**}$. At this point we could assume that $\mathcal{B}$ were reflexive. Then we are guaranteed there exists $\bar{m}_\mu \in \mathcal{B}$ such that

$$m_\mu(l) = l(\bar{m}_\mu) \quad \forall l \in \mathcal{B}^*. \tag{1.6}$$

[1] Seperability of $\mathcal{B}$ is need to ensure f is the pointwise limit of simple functions, i.e. strong measurablity. If we take strong measurability as an additional assumption on f then seperability of $\mathcal{B}$ can be relaxed

However, this is unnecessary supposing $\mathcal{B}$ is seperable. From our first approach of using Bochner integrals, we know in fact that $\int_\mu \|x\| \, \mu(\mathrm{d}x) < \infty$ is a sufficient condition to define $\bar{m}_\mu$ as in. Then since the Bochner integral can be exchanged with any $l \in \mathcal{B}^*$, the property in holds. This shows two things. Firstly, it shows our two approaches recover the same mean. Secondly, it shows reflexivity is an extraneous assumption.

The value of this second approach really shines when the concept we're describing becomes more intangible than a mean, in particular later when we define a Gaussian measure on $\mathcal{B}$

1.2.2. Marginals characterise measures

Seen as we're working with $l(X)$ so much, it's useful to have a notation for its law.

Definition 1.1 (*Pushforward measures*)

Let $(E, \mathcal{E})$ and $(F, \mathcal{F})$ be measurable spaces. Let μ be a measure on $(E, \mathcal{E})$ and $f : (E, \mathcal{E}) \to (F, \mathcal{F})$ be measurable. Then the pushforward *of μ be f is the measure $f_*\mu$ on $(F, \mathcal{F})$ where*

$$f_*\mu(A) \coloneqq \mu(f^{-1}(A)). \tag{1.7}$$

This is defined so that $l(X) \sim l_*\mu$. It is also sensible to have a name for the laws $\{l_*\mu : l \in \mathcal{B}^*\}$.

Definition 1.2 (*One-dimensional marginals*)

Let μ be a probability measure on $\mathcal{B}$. Then the one-dimenional margins *of μ are the laws $l_*\mu$ for all $l \in \mathcal{B}^*$.*

The next proposition shows that the one-dimensional marginals carry enough information to uniquely characterise a meaure μ.

Proposition 1.3 (*Sufficiency of one-dimensional marginals*)

Let μ and ν be Borel probability measures on a seperable Banach space $\mathcal{B}$. Suppose $l_*\mu = l_*\nu$ for all $l \in \mathcal{B}^*$. Then $\mu = \nu$.

Proof. We start in finite dimensions, suppose $\mathcal{B} = \mathbb{R}^n$. Then considering Fourier transforms. All $\xi \in \mathbb{R}^n$ defines a bounded linear functional $l_\xi : \mathbb{R}^n \to \mathbb{R}^n$ by $l_\xi(x) = \xi \cdot x$. Therefore

$$\hat{\mu}(\xi) = \int_{\mathbb{R}^n} e^{i\xi \cdot x} \mu(\mathrm{d}x) = \int_{\mathbb{R}} e^{it} \, (l_\xi)_*\mu(\mathrm{d}t) \tag{1.8}$$

$$= \int_{\mathbb{R}} e^{it} \, (l_\xi)_*\nu(\mathrm{d}t) = \hat{\nu}(\xi). \tag{1.9}$$

Since Fourier transforms uniquely determine probability measures, $\mu = \nu$.

For infinite dimensions, the cylindrical sets of $\mathcal{B}$ are given by

$$\mathrm{Cyl}(\mathcal{B}) \coloneqq \{T^{-1}(A) : T : \mathcal{B} \to \mathbb{R} \text{ linear and bounded}, \ A \subset \mathbb{R}^n \text{ Borel}\}. \tag{1.10}$$

This is a subset of the Borel σ-algebra on $\mathcal{B}$ by continuity of $\mathcal{B}$. This is a π-system since

$$T^{-1}(A) \cap S^{-1}(B) = (T \oplus S)^{-1}(A \times B). \tag{1.11}$$

By the finite dimensional case we just proved, $T_*\mu$ is uniquely determined by

$$\{L_*(T_*\mu) : L \in (\mathbb{R}^n)^*\} = \{(L \circ T)_*\mu : L \in (\mathbb{R}^n)^*\} \subset \{l_*\mu : l \in \mathcal{B}^*\}. \tag{1.12}$$

Thus μ and ν agree on $\mathrm{Cyl}(\mathcal{B})$ and therefore $\sigma(\mathrm{Cyl}(\mathcal{B}))$. The topology on any seperable metric space is generated by countable unions of balls, thus it suffices to show $B_r(x_0) \in \mathrm{Cyl}(\mathcal{B}))$ for any $x_0 \in \mathcal{B}$, $r > 0$. There exists linear functionals $l_n \in \mathcal{B}^*$ such that

$$\|x\| = \sup_n l_n(x). \tag{1.13}$$

Thus

$$B_r(x_0) = \bigcap_n \{x \in \mathcal{B} : l_n(x - x_0) \leq r\} \in \mathrm{Cyl}(\mathcal{B}). \tag{1.14}$$

□

1.2.3. Mean and covariance

We formalise what we said in the introduction.

Definition 1.4 (*Mean Operator*)

Let μ be a probability measure on $\mathcal{B}$ and suppose $l_*\mu$ has finite mean for all $l \in \mathcal{B}^*$. Then the mean operator *of μ is a linear operator $\mu : \mathcal{B}^* \to \mathbb{R}$ where*

$$m_\mu(l) := \int_{\mathcal{B}} l(x)\, \mu(dx). \tag{1.15}$$

μ is centered *if $m_\mu \equiv 0$.*

Similarly we can define the covariance.

Definition 1.5 (*Covariance Operator*)

Let μ be a probability measure on $\mathcal{B}$ and suppose $l_*\mu$ has finite second moment for all $l \in \mathcal{B}^*$. Then the covariance operator *of μ is the bilinear operator $C_\mu : \mathcal{B}^* \times \mathcal{B}^* \to \mathbb{R}$ where*

$$C_\mu(l, l') := \int_{\mathcal{B}} l(x) l'(x) \mu(dx). \tag{1.16}$$

Remark 1.6

Note that C_μ is a symmetric, positive semi-definite, bilinear operator.

We'll leave the question of whether or not these operators are bounded till later.

1.2.4. Fourier transforms

Definition 1.7 (*Fourier Transform*)

The Fourier transform *of a probabilty measure μ on $\mathcal{B}$ is given by $\hat{\mu} \in \mathcal{B}^* \to \mathbb{C}$ where*

$$\hat{\mu}(l) = \int_{\mathcal{B}} e^{il(x)} \mu(dx). \tag{1.17}$$

Remark 1.8

Unlike the mean, the Fourier transform of a measure of $\mathbb{R}^n$ is normally a function to begin with, just a function $\mathbb{R}^n \to \mathbb{R}$ rather than $(\mathbb{R}^n)^* \to \mathbb{R}$. This is a use of the Hilbert space structure on $\mathbb{R}^n$. For a general Hilbert space $\mathcal{H}$, let μ be a probabilty measure on $\mathcal{H}$. We can then instead define $\hat{\mu} : \mathcal{H} \to \mathbb{R}$ by

$$\hat{\mu}(\xi) := \int_{\mathcal{H}} e^{i\langle \xi, x\rangle} \mu(dx). \tag{1.18}$$

This is then equivalent to our definition since by the Riesz representation theorem the map $\xi \to l_\xi$ where $l_\xi(x) = \langle \xi, x\rangle$ is an isomorphism between $\mathcal{H}$ and $(\mathcal{H})^*$.

In the finite dimensional case, Fourier transforms uniquely determine distributions. This holds true in seperable Banach spaces.

Proposition 1.9 (*Uniqueness of Fourier transforms*)

Let μ and ν be probability measures on a seperable Banach space $\mathcal{B}$. If $\hat{\mu} = \hat{\nu}$ then $\mu = \nu$.

Proof. $\hat{\mu}$ uniquely determines the (classical) Fourier transform of $l_*\mu$ since

$$\widehat{l_*\mu}(\xi) = \hat{\mu}(\xi l). \tag{1.19}$$

From the finite dimensional case we know $\widehat{l_*\mu}$ uniquely determines $l_*\mu$. Therefore $\hat{\mu}$ uniquely determines the one-dimensional marginals of μ, and thus μ itself by proposition 1.3 on page 5. □

1.3. Gaussian measures

We've now got our tools ready. Now we'll move onto the next step, defining what a Gaussian random variable is.

1.3.1. Basic definition and mean-covariance characterisation

As aluded to in the section on analysis of $\mathcal{B}$-valued random variables, there is not an immediately obvious definition for what a Gaussian measure is on $\mathcal{B}$ using Bochner

integrals. We can't really hope for a density function after all, there's no generalisation of the Lebesgue measure to general Banach spaces. However we do know what Gaussian measures are on $\mathbb{R}$, so we can use our approach of considering one-dimensional marginals.

Definition 1.10 (*Gaussian measures*)

A Borel probability measure μ on $\mathcal{B}$ is a Gaussian measure *if $l_*\mu$ is a Gaussian measure. Note we consider Dirac measures to be variance 0 Gaussian measures in this definition.*

Gaussian measures on $\mathbb{R}^n$ are uniquely characterised by their covariance and mean. This is true since they uniquely determine the Fourier transform. This remains true in infinite dimensions.

Proposition 1.11 (*Fourier transform of Gaussian measures*)

Let μ be a Gaussian measure on $\mathcal{B}$. Then

$$\hat{\mu}(l) = \exp\left\{ im_\mu(l) - \frac{1}{2}C_\mu(l,l) \right\}. \tag{1.20}$$

Corollary 1.12 (*Mean and cov. uniquely characterise Gaussian measures*)

A Gaussian measure on a seperable Banach space is uniquely determined by it's mean and covariance operators.

1.3.2. Integrability of Gaussian measures

It's hopefully somewhat intuitive that boundedness of the covariance operator will be related to boundedness of the second moment of the corresponding measure. In finite dimensions, Gaussian random variables are nice because of the rapid tail decay of the probability density function. We now show a similar statement is true for general Gaussian measures. Suprisingly, this will come from rotational invariance of Gaussian measure.

Definition 1.13 (*Rotational invariance*)

Let μ be a measure on $\mathcal{B}$. For each $\varphi \in \mathbb{R}$, define $R_\varphi : \mathcal{B} \times \mathcal{B} \to \mathcal{B} \times \mathcal{B}$ by

$$R_\varphi(x,y) = (x\sin\varphi + y\cos\varphi, x\cos\varphi - y\sin\varphi). \tag{1.21}$$

Then μ is invariant under rotations of φ *if $(R_\varphi)_*(\mu \otimes \mu) = \mu \otimes \mu$.*

Proposition 1.14 (*Rotational invariance of centered Gaussians*)

Centered Gaussian measures on a seperable Banach space $\mathcal{B}$ are invariant under rotations by any angle.

Proof. Check the Fourier transform of the rotated measure is equal to the fourier transform of the original measure. □

The following theorem shows as long as a measure is invariance under rotations of $\frac{\pi}{4}$ then it has double exponential tail decay.

Theorem 1.15 (*Fernique*)

Let μ a probability measure on $\mathcal{B}$. Suppose μ is invariant under rotations by $\frac{\pi}{4}$. Then there exists $\alpha > 0$ such that $\|x\|$ has doubly exponential moments, i.e.

$$\int e^{\alpha\|x\|^2}\,\mu(dx) < \infty. \tag{1.22}$$

Proof. Take $t, \tau > 0$. Then

$$\mu(\|x\| \leq \tau)\mu(\|x\| > t) = \mu \otimes \mu\left(\frac{\|x+y\|}{\sqrt{2}} > t, \frac{\|x+y\|}{\sqrt{2}} > \tau\right) \tag{1.23}$$

$$\leq \mu \otimes \mu\left(\|x\| > \frac{t-\tau}{2}, \|x\| > \frac{t-\tau}{2})\right) \tag{1.24}$$

$$= \mu\left(\|x\| > \frac{t-\tau}{2}\right)^2 \tag{1.25}$$

INSERT REST OF PROOF □

Fernique's theorem allows us to not only show all moments of a Gaussian measure are finite, it allows us to bound the higher moments using the first moment.

Proposition 1.16 (*Moment bounds on Gaussians*)

Let μ be a centered Gaussian on $\mathcal{B}$. Then there exists $\alpha, K > 0$ such that for all $n \in \mathbb{N}$

$$\int_{\mathcal{B}} |x|^{2n}\,\mu(dx) \leq n! K \alpha^{-n} M^{2n}. \tag{1.26}$$

where $M = \int_{\mathcal{B}} \|x\|\,\mu(dx)$.

Proof. INSERT PROOF □

1.4. Existence of Gaussian Measures

Existence of Gaussian measures is hard. First let us recall the method in finite dimensions. Given any symmetric, positive semi-definite matrix Σ it is possible to compute its Cholesky factorisation L such that $\Sigma = LL^T$. Then by taking product measures we can construct a standard Gaussian $X \sim N(0, I)$. Then $LX \sim N(0, \Sigma)$.

This section is our first foray into existence results. All our Gaussians will be centered so we just need to worry about the covariance operator. In finite dimensions, we know covariance matrices must be symmetric and positive semi-definite. The first thing we show is that boundedness is required as well when working in infinite dimensions.

Proposition 1.17 (*Boundedness of covariance operators*)

Let μ be a centered Gaussian measure on a seperable Banach space. Then there exists $\|C_\mu\| < \infty$ such that $C_\mu(l, l') \leq \|C_\mu\|\|l\|\|l'\|$. Moreover there exists continuous operator $\hat{C}_\mu : \mathcal{B}^* \to \mathcal{B}$ such that

$$C_\mu(l, l') = l'\left(C_\mu(l)\right). \tag{1.27}$$

Proof. The first part of the proof is clear by setting $\|C_\mu\|$ to $\int_{\mathcal{B}} \|x\|^2 \, \mu(\mathrm{d}x)$ and employing boundedness of l and l'.

For the second part, since

$$\int_{\mathcal{B}} \|xl(x)\| \, \mu(\mathrm{d}x) \leq \|l\| \int_{\mathcal{B}} \|x\|^2 \mathrm{d}x < \infty \tag{1.28}$$

we can formalise

$$\hat{C}_\mu(l) = \int_{\mathcal{B}} xl(x) \, \mu(\mathrm{d}x) \tag{1.29}$$

as a Bochner integral. Then since the Bochner integral can be exchanged with continuous linear functionals

$$\begin{aligned} l'(\hat{C}_\mu(l)) &= l'\left(\int_{\mathcal{B}} xl(x) \, \mu(\mathrm{d}x)\right) = \int_{\mathcal{B}} l'(xl(x)) \, \mu(\mathrm{d}x) \\ &= \int_{\mathcal{B}} l(x)l'(x)\mu(\mathrm{d}x) = C_\mu(l, l') \end{aligned} \tag{1.30}$$

as required. □

So now let's add boundedness to our list of assumptions. Is that enough to characterize covariance operators of Gaussian measures? Unfortunately the answer is still no. In fact for a general Banach space there is no nice characterisation. However for Hilbert spaces we can give an actual answer. As promised, we now characterise the covariance operators of Gaussian measures in Hilbert spaces. In proposition 1.17 we've already shown there exists $\hat{C}_\mu : \mathcal{H}^* \to \mathcal{H}$ such that $C_\mu(l, l') = l'(\hat{C}_\mu(l))$ for all $l, l' \in \mathcal{B}^*$. Then by the Riesz representation, there exists $\tilde{C}_\mu : \mathcal{H} \to \mathcal{H}$ such that

$$C_\mu(\langle \cdot, x\rangle, \langle \cdot, y\rangle) = \langle \tilde{C}_\mu x, y\rangle \tag{1.31}$$

We will state our characterisation in terms of $\tilde{C}_\mu$. Firstly let's recall some definitions from functional analysis.

Definition 1.18 (*Positive definiteness and trace-class*)

Let $\mathcal{H}$ be a seperable Hilbert space and let $L : \mathcal{H} \to \mathcal{H}$ be linear.

1. *L is* positive definite *if L is self-adjoint (symmetric) and $\langle Lh, h\rangle > 0$ for all non-zero $h \in \mathcal{H}$*

2. *L is* trace-class *if for any orthonormal basis (e_n) of $\mathcal{H}$ there exists a sequence of numbers $\{\lambda_n\}$ such that $\sum_n |\lambda_n| < \infty$ and*

$$\langle x, Lx\rangle = \sum_n \lambda_n \langle x, e_n\rangle^2. \tag{1.32}$$

The trace *of L is then $\operatorname{Tr} L := \sum_n \lambda_n$.*

You can prove it is sufficient to check the trace-class property for any single orthonormal basis in order for it to hold for all orthonormal bases.

Trace-class is a stronger property than boundedness, in particular there exists operators that are bounded and not trace class. Hence we see symmetry, positive semi-definitess, and boundedess was truly insufficient to characterize Gaussian covariances.

Proposition 1.19 (*Characterisation of Gaussian covariances in Hilbert Spaces*)

Let $\mathcal{H}$ be a seperable Hilbert space and $\tilde{C} : \mathcal{H} \to \mathcal{H}$. Then the following are equivalent

1. *There exists a Gaussian measure μ on $\mathcal{H}$ such that*

$$C_\mu(\langle \cdot, x\rangle, \langle \cdot, y\rangle) = \langle \tilde{C}x, y\rangle \quad \forall x, y \in \mathcal{H}. \tag{1.33}$$

2. *$\tilde{C}$ is a linear operator that is positive semi-definite and trace-class.*

Proof. (1) $\implies$ (2): Bilinerity, symmetry and positive semi-definiteness of $\hat{C}_\mu$ translates to linearity, self-adjointness and positive semi-definiteness of $\tilde{C}$ respectively. To show trace-class fix any orthonormal basis $\{e_n\}$. Then

$$\begin{aligned}\operatorname{Tr}(\tilde{C}) &= \sum_n \langle \tilde{C}e_n, e_n\rangle = \sum_n C_\mu(\langle \cdot, e_n\rangle, \langle \cdot, e_n\rangle) = \sum_n \int_{\mathcal{H}} \langle x, e_n\rangle^2 \, \mu(\mathrm{d}x) \\ &= \int_{\mathcal{H}} \sum_n \langle x, e_n\rangle^2 \, \mu(\mathrm{d}x) = \int_{\mathcal{H}} \|x\|^2 \, \mu(\mathrm{d}x) < \infty \end{aligned} \tag{1.34}$$

using proposition 1.16 on page 9.

(2) $\implies$ (1): Trace-class operators are compact. The spectral theorem for compact, self-adjoint operators says there is an orthonormal basis $\{e_n\}$ of $\mathcal{H}$ consisting of eigenvectors of $\tilde{C}$, say $\tilde{C}(e_n) = \lambda_n e_n$, and that the λ_n are real. Furthermore $\lambda_n \geq 0$ for all n

as $\tilde{C}$ is positive semi-definite. In addition

$$\sum_n \lambda_n = \sum_n \langle \tilde{C} e_n, e_n \rangle = \mathrm{Tr}(\tilde{C}) < \infty. \tag{1.35}$$

There exists a probability space on which we can define $X_1, X_2, \ldots$ a sequence of i.i.d. $N(0,1)$ random variables. Then let μ_n be the law of $\sum_{i=1}^n \sqrt{\lambda_i} X_i e_i$. We can show μ_n converges weakly to a limit μ and that μ is a Gaussian measure with the required covariance. $\square$

1.4.1. Cameron-Martin theory

A. Miscellaneous Theorems

Proposition A.1 (*Seperable metric spaces have a countable base of balls*)

Let (X, d) be a seperable metric space and $A \subset X$ be open. Let D be a countable dense set of X. Then

$$A = \bigcup_{x \in D \cap A} \bigcup_{\substack{q \in \mathbb{Q} \\ B_q(x) \subset A}} B_q(x). \tag{A.1}$$

In other words $\{B_q(x) : q \in \mathbb{Q}, x \in D\}$ is a countable base of the topology on (X, d).

Proof. Fix any $y \in A$. Then by openess there exists $r > 0$ such that $B_r(y) \subset A$. By density there exists $x \in B_{r/4}(y) \cap D$. Pick rational $q \in \left(\frac{1}{4}r, \frac{3}{4}r\right)$. Then $y \in B_q(x)$ and by triangle inequality $B_q(x) \subset B_{\frac{r}{4}+q}(y) \subset B_r(y) \subset A$. □

Proposition A.2 (*Approximation of norm in seperable spaces*)

Suppose $\mathcal{B}$ is a seperable Banach space. Then there exists countable many linear functions $l_n \in \mathcal{B}^*$ such that

$$\|x\| = \sup_n l_n(x) \quad \forall x \in \mathcal{B}. \tag{A.2}$$

Proof. Let $\mathcal{D} = \{x_n\}$ be a countable dense set of $\mathcal{B}$. Then by Hahn-Banach, for each x_n there exists $l_n \in \mathcal{B}^*$ such that $\|l_n\| \leq 1$ and $l_n(x_n) = \|x_n\|$. Since $\|l_n\| \leq 1$

$$\|x\| \geq \sup_n l_n(x) \quad \forall x \in \mathcal{B}. \tag{A.3}$$

To show equality, for any $x \in \mathcal{B}$ take a sequence $\{x_n\} \subset \mathcal{D}$ such that $x_n \to x$ in $\mathcal{B}$. Then

$$\begin{aligned}
|l_n(x) - \|x\|| &= |l_n(x) - \|x_n\| + \|x_n\| - \|x\|| && \text{(A.4)}\\
&= |l_n(x) - l_n(x_n) + \|x_n\| - \|x\|| && \text{(A.5)}\\
&\leq |l_n(x - x_n)| + |\|x_n\| - \|x\|| && \text{(A.6)}\\
&\leq \|x - x_n\| + |\|x_n\| - \|x\|| \xrightarrow{n\to\infty} 0. && \text{(A.7)}
\end{aligned}$$

Therefore $l_n(x) \to \|x\|$. Thus we have equality. □

Proposition A.3 (*Almost sure convergence implies weak convergence of laws*)

Let μ_n, μ be probability measures on a metric space S. Suppose there exists random

variables $X_n, X : \Omega \to S$ defined on a common probability space such that $X_n \sim \mu_n$, $X \sim \mu$, and $X_n \to X$ almost surely. Then $\mu_n \Rightarrow \mu$.

Proof. Let $f : S \to \mathbb{R}$ be bounded and continuous. Then

$$\int_S f(s)\,\mu_n(\mathrm{d}x) = \mathbb{E}[f(X_n)] \to E[f(X)] = \int_S f(x)\,\mu(\mathrm{d}x) \tag{A.8}$$

as $n \to \infty$ by the dominated convergence theorem. Thus $\mu_n \Rightarrow \mu$. □

The converse is also true assuming μ has seperable support, that if we have weak convergence we can find representations which converge almost surely. This is the Skorokhod representation theorem.

*

Proof. Suppose $\mathcal{A}$ were not dense in $C([0,1]^d, \mathbb{R})^*$, say l_0 is not in the weak-* closure of $\mathcal{A}$. Then by the Hahn-Banach seperation, there exists a weak-* continuous linear functional $L : C([0,1]^d, \mathbb{R})^* \to \mathbb{R}$ that seperates the closed convex set $\bar{\mathcal{A}}$ and the compact convex set $\{l_0\}$, i.e. there exists $\alpha \in \mathbb{R}$ such that

$$L(l) \leq \alpha \leq L(l_0) \quad \forall l \in \bar{\mathcal{A}}. \tag{A.9}$$

Then as $\mathcal{A}$ is a vector space, for all $l \in \mathcal{A}$ and $\lambda \in \mathbb{R}$ we have $\lambda L(l) = L(\lambda l) \leq \alpha$. Therefore $L(l) = 0$. Since L is continuous w.r.t. the weak-$*$ topology there exists $f \in C([0,1]^d, \mathbb{R})$ such that $L(l) = l(f)$ for all $l \in \mathcal{B}^*$. Then $f(x) = L(\delta_x) = 0$ for all $x \in \mathcal{D}$. Thus $f \equiv 0$ by density of $\mathcal{D}$ and continuity of f. Hence $L \equiv 0$, contradiction. □

When we use this in the proof of KCC we actually have $\mathcal{D} = [0,1]^d$ rendering density trivial.

Proposition A.4 (*Gaussians converge weakly to Gaussians*)

Let μ_n be a sequence of Gaussian laws on $\mathcal{B}$. Suppose $\mu_n \Rightarrow \mu$. Then μ is Gaussian.

Proof. Start with $\mathcal{B} = \mathbb{R}$, suppose each $\mu_n \sim N(\theta_n, \sigma_n)$. (μ_n) is a tight sequence since it is weakly convergent. By tightness there exists an bounded interval $[K, -K]$ such that for all n, $\mu_n([-K, K]) > \frac{1}{2}$. Then $\theta_n \in [-K, K]$ for all n else over half the mass of some μ_n would lie outside $[-K, K]$. Similarly

$$\mu_n([-K, K]) = \int_{-K}^{K} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(x-\theta_n)^2} \mathrm{d}x \leq K\sqrt{\frac{2}{\pi\sigma_n^2}}. \tag{A.10}$$

Thus (σ_n) must be bounded else the above quantity could be reduced below $\frac{1}{2}$. Hence there exist a subsequence along which $\theta = \lim_n \theta_{k_n}$ and $\sigma = \lim_n \sigma_{k_n}$ both exist. Then since $x \mapsto e^{i\xi x}$ is a continuous bounded function, by definition of weak convergence

$$\hat{\mu}(\xi) = \lim_n \hat{\mu}_n(\xi) = \lim_n \hat{\mu}_{k_n}(\xi) = \lim_n e^{i\theta_n\xi - \frac{1}{2}\sigma_n^2\xi^2} = e^{i\theta\xi - \frac{1}{2}\sigma^2}. \tag{A.11}$$

Since Fourier transforms uniquely determine probability measures, $\mu \sim N(\theta, \sigma^2)$.

For general $\mathcal{B}$, if $\mu_n \Rightarrow \mu$ then $l_*\mu_n \to l_*\mu$ for all $l \in \mathcal{B}^*$. Therefore $l_*\mu$ is Gaussian by the $\mathcal{B} = \mathbb{R}$ case. Hence μ is Gaussian. □

Proposition A.5 (*Mutual Singularity and Total Variation*)

Let μ and ν be probability measures on the measurable space $(E, \mathcal{E})$. Recall the total variation

$$\|\mu - \nu\|_{TV} = \sup_{A \in \mathcal{E}} |\mu(A) - \nu(A)|. \tag{A.12}$$

Then the following are equivalent:

1. *μ and ν are mutually singular.*
2. *$\|\mu - \nu\|_{TV} = 1$.*

Proof. (1) $\implies$ (2) is clear. For the converse if $\|\mu - \nu\|_{TV} = 1$ then there exists $\{A_n\} \subset \mathcal{E}$ such that

$$\lim_n |\mu(A_n) - \nu(A_n)| = 1 \tag{A.13}$$

The sequence $(\mu(A_n))$ and $(\nu(A_n))$ are bounded so by Bolzano-Weierstrass we can pass to a subsequence along which $\alpha := \lim_n \mu(A_n)$ and $\beta := \lim_n \nu(A_n)$ both exist. Then $\alpha, \beta \in [0, 1]$ but also $|\alpha - \beta| = 1$. The only way this holds is if $\alpha = 0$ and $\beta = 1$ or $\alpha = 1$ and $\beta = 0$. WLOG we have the case $\alpha = 0$ and $\beta = 1$. In other words $\lim_n \mu(A_n) = 0$ so by passing to a further subsequence we can WLOG that

$$\mu(A_n) \leq 2^{-n}. \tag{A.14}$$

Let $A = \{A_n \text{ occurs infinitely often}\} = \cap_{n \geq 1} \cup_{k \geq n} A_k$. Then $\sum_n \mu(A_n) < \infty$ thus $\mu(A) = 0$ by the first Borel Cantelli lemma. Finally

$$\nu(A) = \lim_n \nu(\cup_{k \geq n} A_k) \geq \lim_n \nu(A_n) = 1. \tag{A.15}$$

Thus $\mu(A) = 0$ and $\nu(A) = 1$, showing that μ and ν are mutually singular. □