

# The narrow topology on the set of Borel probability measures on a metrizable space

Jordan Bell

jordan.bell@gmail.com

Department of Mathematics, University of Toronto

March 29, 2015

## 1 Introduction

In this note we talk about the total variation metric and the narrow topology (also called the weak-\* topology) on the set of Borel probability measures on a metrizable topological space. In the course of the note we define all the terms used in the previous sentence.

The reason for talking about metrizable spaces rather than metric spaces in this note is to make explicit that the objects we define do not depend on a metric but only on the topological properties of a space. On the other hand, the Prokhorov metric, which we do not talk about in this note, depends on the metric, not merely on the fact that a topological space is metrizable.

## 2 Preliminaries

If  $\mathfrak{M}$  is a  $\sigma$ -algebra on a set  $X$ , a **positive measure** is a function  $\mu : \mathfrak{M} \rightarrow [0, \infty]$  such that  $\{A_i\}$  being a countable subset of  $\mathfrak{M}$  with pairwise disjoint members implies that

$$\mu\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mu(A_i).$$

A positive measure  $\mu$  is called **finite** if  $X$  has finite measure, a **probability measure** if  $\mu(X) = 1$ , and  **$\sigma$ -finite** if there is a countable collection  $\{E_i\} \subset \mathfrak{M}$  each member of which has finite measure and that satisfies  $X = \bigcup_{i=1}^{\infty} E_i$ .

Suppose that  $\mathfrak{M}$  is a  $\sigma$ -algebra on a set and that  $E \in \mathfrak{M}$ . A countable collection  $\{E_i\} \subset \mathfrak{M}$  whose members are pairwise disjoint and that satisfies  $E = \bigcup_{i=1}^{\infty} E_i$  is called a **partition of  $E$** . A **signed measure** is a function  $\mu : \mathfrak{M} \rightarrow \mathbb{R}$  such that if  $E \in \mathfrak{M}$ , then

$$\mu(E) = \sum_{i=1}^{\infty} \mu(E_i) \tag{1}$$

for every partition  $\{E_i\}$  of  $E$ .

Any statement we make about a measure on a  $\sigma$ -algebra without specifying whether it is a positive measure or a signed measure applies to both classes.

### 3 Total variation

If  $\mathfrak{M}$  is a  $\sigma$ -algebra on a set  $X$  and  $\mu$  is a signed measure on  $\mathfrak{M}$ , for  $E \in \mathfrak{M}$  we define

$$|\mu|(E) = \sup \left\{ \sum_{i=1}^{\infty} |\mu(E_i)| : \{E_i\} \text{ is a partition of } E \right\}.$$

We call  $|\mu|$  the **total variation of  $\mu$** . One proves that if  $\mu$  is a signed measure on  $\mathfrak{M}$ , then the total variation  $|\mu|$  is a finite positive measure on  $\mathfrak{M}$ .<sup>1</sup>

If  $\mathfrak{M}$  is a  $\sigma$ -algebra on a set, we denote by  $ca(\mathfrak{M})$  the set of signed measures on  $\mathfrak{M}$ ; the notation  $ca$  stands for “countably additive”. For  $\mu, \lambda \in ca(\mathfrak{M})$  and  $c \in \mathbb{R}$ , we define

$$(a\mu + \lambda)(E) = a\mu(E) + \lambda(E), \quad E \in \mathfrak{M}.$$

It is straightforward to check that  $a\mu + \lambda \in ca(\mathfrak{M})$ , so that  $ca(\mathfrak{M})$  is a vector space. We check that

$$\|\mu\| = |\mu|(X),$$

is a norm  $ca(\mathfrak{M})$ , called the **total variation norm**.

**Theorem 1.** *Suppose that  $\mathfrak{M}$  is a  $\sigma$ -algebra on a set. Then  $ca(\mathfrak{M})$  is a real Banach space with the total variation norm.*

Suppose that  $\mu$  is a positive measure on a  $\sigma$ -algebra  $\mathfrak{M}$  and that  $\lambda$  is a measure on  $\mathfrak{M}$ . We say that  $\lambda$  **absolutely continuous with respect to  $\mu$** , denoted

$$\lambda \ll \mu,$$

if  $E \in \mathfrak{M}$  and  $\mu(E) = 0$  imply that  $\lambda(E) = 0$ .

If  $A \in \mathfrak{M}$  and  $\lambda(E) = \lambda(A \cap E)$  for every  $E \in \mathfrak{M}$ , we say that  $\lambda$  is **concentrated on  $A$** . If  $\lambda_1, \lambda_2$  are measures on  $\mathfrak{M}$  and there are disjoint sets  $A_1, A_2 \in \mathfrak{M}$  such that  $\lambda_1$  is concentrated on  $A_1$  and  $\lambda_2$  is concentrated on  $A_2$ , we say that  $\lambda_1$  and  $\lambda_2$  are **mutually singular** and write

$$\lambda_1 \perp \lambda_2.$$

The following theorem states both the **Lebesgue decomposition** and the **Radon-Nikodym theorem**.<sup>2</sup>

<sup>1</sup>Walter Rudin, *Real and Complex Analysis*, third ed., pp. 117-118, Theorem 6.2 and Theorem 6.4.

<sup>2</sup>Walter Rudin, *Real and Complex Analysis*, third ed., p. 121, Theorem 6.10.

**Theorem 2** (The Lebesgue decomposition and the Radon-Nikodym theorem). *Suppose that  $\mathfrak{M}$  is a  $\sigma$ -algebra on a set  $X$  and that  $\mu$  is a  $\sigma$ -finite positive measure on  $\mathfrak{M}$ . If  $\lambda \in ca(\mathfrak{M})$ , then:*

1. *There is a unique pair  $(\lambda_a, \lambda_s)$  of elements of  $ca(\mathfrak{M})$  such that*

$$\lambda = \lambda_a + \lambda_s, \quad \lambda_a \ll \mu, \quad \lambda_s \perp \mu.$$

*If  $\lambda$  is a finite positive measure, then so are  $\lambda_a$  and  $\lambda_s$ .*

2. *There is a unique  $h \in L^1(\mu)$  such that*

$$\lambda_a(E) = \int_E h d\mu, \quad E \in \mathfrak{M}.$$

The function  $h \in L^1(\mu)$  in the above theorem is called the **Radon-Nikodym derivative of  $\lambda$  with respect to  $\mu$** , and we write

$$d\lambda = h d\mu.$$

Suppose that  $\mu$  is a signed measure on a  $\sigma$ -algebra  $\mathfrak{M}$ . Then  $|\mu|$  is a finite positive measure on  $\mathfrak{M}$ . For any  $E \in \mathfrak{M}$ ,  $|\mu|(E) \geq |\mu(E)|$ , so  $\mu$  is absolutely continuous with respect to  $|\mu|$  and by the Radon-Nikodym theorem there is some  $h \in L^1(|\mu|)$  such that

$$d\mu = h d|\mu|.$$

One proves<sup>3</sup> that  $|h| = 1$ , that is, that for  $|\mu|$ -almost all  $x$ ,  $h(x) \in \{-1, 1\}$ . Using this function  $h$ , we define

$$\int f d\mu = \int f h d|\mu|.$$

The right-hand side is the integral of a real valued function with respect to a finite positive measure.

If one wishes to speak only about probability measures rather than signed measures, one might choose to use the expression in the following lemma to define a metric on the set of probability measures.

**Lemma 3.** *If  $\mu$  and  $\nu$  are probability measures on a  $\sigma$ -algebra  $\mathfrak{M}$  on a set  $X$ , then*

$$\|\mu - \nu\| = 2 \sup_{E \in \mathfrak{M}} |\mu(E) - \nu(E)|.$$

*Proof.* Let  $m = \mu + \nu$ . Then  $\mu$  and  $\nu$  are each absolutely continuous with respect to  $m$ , so by the Radon-Nikodym theorem there are  $f, g \in L^1(m)$  such that  $d\mu = f dm$  and  $d\nu = g dm$ . Let  $\lambda = \mu - \nu$ , which satisfies  $d\lambda = (f - g) dm$ , and write  $h = f - g$ . For any  $E \in \mathfrak{M}$  we have

$$\lambda(E) + \lambda(X \setminus E) = \lambda(X) = \mu(X) - \nu(X) = 0,$$

---

<sup>3</sup>Walter Rudin, *Real and Complex Analysis*, third ed., p. 124, Theorem 6.12.

hence  $|\lambda(E)| = |\lambda(X \setminus E)|$  and therefore

$$\begin{aligned} |\lambda(E)| &= \frac{1}{2}(|\lambda(E)| + |\lambda(X \setminus E)|) \\ &= \frac{1}{2} \left( \left| \int_E h dm \right| + \left| \int_{X \setminus E} h dm \right| \right) \\ &\leq \frac{1}{2} \left( \int_E |h| dm + \int_{X \setminus E} |h| dm \right) \\ &= \frac{1}{2} \int_X |h| dm. \end{aligned}$$

This shows that

$$\sup_{E \in \mathfrak{M}} |\lambda(E)| \leq \frac{1}{2} \int_X |h| dm.$$

Let  $E = \{x \in X : h(x) > 0\} \in \mathfrak{M}$ . For this set  $E$ ,

$$\begin{aligned} |\lambda(E)| &= \frac{1}{2} \left( \left| \int_E h dm \right| + \left| \int_{X \setminus E} h dm \right| \right) \\ &= \frac{1}{2} \left( \int_E h dm - \int_{X \setminus E} h dm \right) \\ &= \frac{1}{2} \left( \int_E |h| dm + \int_{X \setminus E} |h| dm \right) \\ &= \frac{1}{2} \int_X |h| dm. \end{aligned}$$

Therefore the previous inequality is in fact an equality:

$$\sup_{E \in \mathfrak{M}} |\lambda(E)| = \frac{1}{2} \int_X |h| dm.$$

But  $d\lambda = h dm$  implies that

$$\|\lambda\| = \int_X |h| dm;$$

this is proved in Rudin.<sup>4</sup> Therefore

$$\|\mu - \nu\| = \|\lambda\| = 2 \sup_{E \in \mathfrak{M}} |\lambda(E)| = 2 \sup_{E \in \mathfrak{M}} |\mu(E) - \nu(E)|.$$

□

---

<sup>4</sup>Walter Rudin, *Real and Complex Analysis*, third ed., p. 125, Theorem 6.13.

## 4 Borel measures

Suppose that  $X$  is a topological space. The smallest  $\sigma$ -algebra containing the topology of  $X$  is called the **Borel  $\sigma$ -algebra of  $X$** . The Borel  $\sigma$ -algebra of  $X$  is denoted  $\mathcal{B}_X$  and elements of  $\mathcal{B}_X$  are called **Borel sets**. A measure on  $\mathcal{B}_X$  is called a **Borel measure on  $X$** .

Suppose that  $\mu$  is a positive Borel measure on  $X$ . We say that  $\mu$  is **inner regular** if for every Borel set  $E$  we have

$$\mu(E) = \sup\{\mu(C) : C \subset E \text{ and } C \text{ is closed}\}$$

and **outer regular** if for every Borel set  $E$  we have

$$\mu(E) = \inf\{\mu(V) : E \subset V \text{ and } V \text{ is open}\}.$$

If  $X$  is Hausdorff, we say that  $\mu$  is **tight** if for every  $\epsilon > 0$  there is some compact subset  $K$  of  $X$  such that  $\mu(X \setminus K) < \epsilon$ . It is straightforward to check that a finite positive Borel measure is tight if and only if for every Borel set  $E$ ,

$$\mu(E) = \sup\{\mu(K) : K \subset E \text{ and } K \text{ is compact}\}.$$

We specify that  $X$  be Hausdorff to ensure that any compact subset of  $X$  is closed and hence a Borel set.

It can be proved that if  $X$  is a metrizable space, then any finite positive Borel measure on  $X$  is inner regular and outer regular,<sup>5</sup> and that if  $X$  is a Polish space then any finite positive Borel measure on  $X$  is tight.<sup>6</sup> (A **Polish space** is a separable topological space whose topology is induced by a complete metric.)

For a topological space  $X$ , we define  $C(X)$  to be the set of continuous functions  $X \rightarrow \mathbb{R}$  and  $C_b(X)$  to be the set of bounded continuous functions  $X \rightarrow \mathbb{R}$ . One checks that  $C_b(X)$  with the supremum norm is a Banach space. We remind ourselves that if  $X$  is compact then  $C(X) = C_b(X)$ .

For a topological space  $X$ , we write  $ca(X) = ca(\mathcal{B}_X)$ . For a compact metrizable space, the **Riesz representation theorem** states that  $\Lambda : ca(X) \rightarrow C(X)^*$  defined by

$$\Lambda_\mu f = \int_X f d\mu, \quad \mu \in ca(X), \quad f \in C(X),$$

is an isomorphism of real Banach spaces, where  $C(X)^*$  has the operator norm.<sup>7</sup>

## 5 The narrow topology on $\mathcal{P}(X)$

Suppose that  $X$  is a topological space. The **narrow topology** on  $ca(X)$  is the coarsest topology such that for each  $f \in C_b(X)$ , the map  $\mu \mapsto \int_X f d\mu$

<sup>5</sup>Alexander S. Kechris, *Classical Descriptive Set Theory*, p. 107, Theorem 17.10.

<sup>6</sup>Alexander S. Kechris, *Classical Descriptive Set Theory*, p. 107, Theorem 17.11.

<sup>7</sup>Charalambos D. Aliprantis and Kim C. Border, *Infinite Dimensional Analysis: A Hitchhiker's Guide*, third ed., p. 498, Corollary 14.15.

is continuous  $ca(X) \rightarrow \mathbb{R}$ . The narrow topology is also called the **weak-\* topology**.

Suppose that  $X$  is a metrizable topological space. We denote by  $\mathcal{P}(X)$  the set of Borel probability measures on  $X$ .  $\mathcal{P}(X)$  is a convex subset of  $ca(X)$ . We shall be interested in the narrow topology on  $\mathcal{P}(X)$ . This is the coarsest topology such that for each  $f \in C_b(X)$ , the map  $\mu \mapsto \int_X f d\mu$  is continuous  $\mathcal{P}(X) \rightarrow \mathbb{R}$ , or equivalently, the subspace topology on  $\mathcal{P}(X)$  inherited from  $ca(X)$  with the narrow topology.

**Lemma 4.** *Suppose that  $X$  is a metrizable topological space. Then  $\mathcal{P}(X)$  is a closed subset of  $ca(X)$  with the narrow topology.*

It is a fact that if  $\mu, \nu$  are Borel probability measures on a metrizable topological space  $X$ , then  $\mu = \nu$  if and only if  $\int_X f d\mu = \int_X f d\nu$  for all  $f \in C_b(X)$ .<sup>8</sup> Thus, the map  $\mathcal{P}(X) \rightarrow \mathbb{R}^{C_b(X)}$  defined by  $\mu \mapsto (f \mapsto \int_X f d\mu)$  is one-to-one.

A useful characterization of the narrow topology on  $\mathcal{P}(X)$  is the following, called the **portmanteau theorem**.<sup>9</sup>

**Theorem 5.** *Suppose that  $X$  is a metrizable topological space, that  $\mu \in \mathcal{P}(X)$ , and that  $\mu_i$  is a net in  $\mathcal{P}(X)$ . The net  $\mu_i$  narrowly converges to  $\mu$  if and only if for every closed set  $C$  in  $X$ ,  $\limsup_i \mu_i(C) \leq \mu(C)$ .*

For  $x \in X$ , define  $\delta_x$  on  $\mathcal{B}_X$  by

$$\delta_x(E) = \begin{cases} 1 & x \in E \\ 0 & x \notin E, \end{cases} \quad E \in \mathcal{B}_X.$$

Then  $\delta_x \in \mathcal{P}(X)$ . We prove that the mapping  $x \mapsto \delta_x$  is an embedding of  $X$  into  $\mathcal{P}(X)$  with the narrow topology, and that if  $X$  is separable then its image is closed.<sup>10</sup>

**Theorem 6.** *Suppose that  $X$  is a metrizable topological space and assign  $\mathcal{P}(X)$  the narrow topology. Then  $x \mapsto \delta_x$  is a homeomorphism onto its image, and if  $X$  is separable then the image is a closed subset of  $\mathcal{P}(X)$ .*

*Proof.* It is a fact that because  $X$  is metrizable, a net  $x_i$  converges to  $x$  if and only if  $f(x_i) \rightarrow f(x)$  for every  $f \in C_b(X)$ ; for these to be equivalent, it suffices that  $X$  be a completely regular topological space.<sup>11</sup>  $f(x_i) \rightarrow f(x)$  can be written

$$\int_X f d\delta_{x_i} \rightarrow \int_X f d\delta_x.$$

<sup>8</sup>Charalambos D. Aliprantis and Kim C. Border, *Infinite Dimensional Analysis: A Hitchhiker's Guide*, third ed., p. 506, Theorem 15.1.

<sup>9</sup>Charalambos D. Aliprantis and Kim C. Border, *Infinite Dimensional Analysis: A Hitchhiker's Guide*, third ed., p. 508, Theorem 15.3.

<sup>10</sup>Charalambos D. Aliprantis and Kim C. Border, *Infinite Dimensional Analysis: A Hitchhiker's Guide*, third ed., p. 512, Theorem 15.8.

<sup>11</sup>Charalambos D. Aliprantis and Kim C. Border, *Infinite Dimensional Analysis: A Hitchhiker's Guide*, third ed., p. 50, Corollary 2.57. A topological space is called **completely regular** if any closed set and any point not in this closed set can be separated by a continuous function.

This being true for all  $f \in C_b(X)$  means that  $\delta_{x_i} \rightarrow \delta_x$ . This shows that  $x \mapsto \delta_x$  is a homeomorphism between  $X$  and  $\{\delta_x : x \in X\}$ .

We remind ourselves that the **support** of a positive Borel measure  $\mu$  on a topological space  $(Y, \tau)$  is the set of all  $y \in Y$  such that if  $N$  is an open neighborhood of  $y$  then  $\mu(N) > 0$ . The support of  $\mu$  is denoted  $\text{supp } \mu$ , and can be written

$$\text{supp } \mu = X \setminus \bigcup_{V \in \tau, \mu(V)=0} V,$$

which makes it apparent that  $\text{supp } \mu$  is a closed set. One checks that  $G \in \tau$  and  $G \cap \text{supp } \mu \neq \emptyset$  imply that  $\mu(G \cap \text{supp } \mu) > 0$ ;  $\text{supp } \mu$  is in fact the largest closed set with this property. One can prove that if  $Y$  is second-countable, then  $\mu(X \setminus \text{supp } \mu) = 0$ .<sup>12</sup>

Suppose that  $X$  is separable and suppose that  $\delta_{x_i} \rightarrow \mu$  in  $\mathcal{P}(X)$ . Since  $X$  is separable and metrizable it is second-countable, so  $\mu(X \setminus \text{supp } \mu) = 0$ , which because  $\mu$  is a finite measure can be written  $\mu(\text{supp } \mu) = \mu(X) = 1$ . Let  $x \in \text{supp } \mu$  and let  $N$  be an open neighborhood of  $x$ . A metrizable space is completely regular, so there is some continuous function  $f : X \rightarrow [0, 1]$  such that  $f(x) = 1$  and  $f(y) = 0$  for all  $y \in X \setminus N$ . This function  $f$  belongs to  $C_b(X)$  as its range is contained in  $[0, 1]$ . Let  $G = \{w \in X : f(w) > \frac{1}{2}\}$ , which is an open set containing  $x$  and hence  $\mu(G) > 0$ . Then,

$$\int_X f d\mu \geq \int_G f d\mu \geq \frac{1}{2} \int_G d\mu = \frac{1}{2} \mu(G) > 0.$$

$\delta_{x_i} \rightarrow \mu$  implies

$$f(x_i) = \int_X f d\delta_{x_i} \rightarrow \int_X f d\mu.$$

Thus  $f(x_i)$  has a positive limit, and so there is some  $i_0$  such that  $i \geq i_0$  implies that  $f(x_i) > 0$ . But from the definition of  $f$ , if  $y \in X \setminus N$  then  $f(y) = 0$ , so  $f(y) > 0$  implies that  $y \in N$ . Thus  $x_i \in N$  for  $i \geq i_0$ . This shows that for every open neighborhood  $N$  of  $x$  there is some  $i_0$  such that  $i \geq i_0$  implies that  $x_i \in N$ , which is what it means to say that  $x_i \rightarrow x$ . Hence, if  $x \in \text{supp } \mu$  then the net  $x_i$  converges to  $x$ . But  $X$  is Hausdorff and we know that  $\text{supp } \mu \neq \emptyset$ , so  $\text{supp } \mu$  has exactly one element. Write  $\text{supp } \mu = \{x\}$ , and check that  $\mu = \delta_x$ , which completes the proof.  $\square$

One can further prove that if  $X$  is a separable metrizable topological space, then<sup>13</sup>

$$\text{ext } \mathcal{P}(X) = \{\delta_x : x \in X\}.$$

We remind ourselves that if  $C$  is a subset of a vector space  $V$  then  $x \in V$  is called an **extreme point** of  $C$  if  $x = (1-t)y + tz$  with  $0 < t < 1$  implies that

<sup>12</sup>Charalambos D. Aliprantis and Kim C. Border, *Infinite Dimensional Analysis: A Hitchhiker's Guide*, third ed., p. 442, Theorem 12.14.

<sup>13</sup>Charalambos D. Aliprantis and Kim C. Border, *Infinite Dimensional Analysis: A Hitchhiker's Guide*, third ed., p. 512, Theorem 15.19.

$y = x$  and  $z = x$ , and  $\text{ext } C$  denotes the set of extreme points of  $C$ . Here the vector space is  $ca(X)$ , of which  $\mathcal{P}(X)$  is a convex subset.

We only speak about  $\mathcal{P}(X)$  when  $X$  is a metrizable topological space, but  $\mathcal{P}(X)$  need not itself be metrizable. However, the following theorem shows that  $X$  is compact if and only if  $\mathcal{P}(X)$  is both compact and metrizable.<sup>14</sup>

**Theorem 7.** *Suppose that  $X$  is a metrizable topological space. Then  $X$  is compact if and only if  $\mathcal{P}(X)$  with the narrow topology is compact and metrizable.*

*Proof.* Suppose that  $X$  is compact. Then  $C(X)$  is a separable Banach space. Denote the closed unit ball in  $C(X)^*$  by  $B$  and assign  $B$  the subspace topology inherited from  $C(X)^*$  with the weak-\* topology. Because  $C(X)$  is a normed space the Banach-Alaoglu theorem tells us that  $B$  is compact, and because  $C(X)$  is a separable normed space,  $B$  is metrizable.<sup>15</sup> Earlier in this note we gave a statement of the Riesz representation theorem for the case of compact metrizable topological spaces: the map  $\Lambda : ca(X) \rightarrow C(X)^*$  defined by

$$\Lambda_\mu f = \int_X f d\mu, \quad \mu \in ca(X), \quad f \in C(X),$$

is an isomorphism of real Banach spaces. Check that  $\Lambda\mathcal{P}(X)$  is a closed subset of  $B$  and hence is compact, and therefore is a weak-\* compact subset of  $C(X)^*$ . Because  $B$  is metrizable and  $\Lambda\mathcal{P}(X)$  is a subset of  $B$ , the subspace topology on  $\Lambda\mathcal{P}(X)$  inherited from  $B$  is metrizable, and this topology is the same as the subspace topology on  $\Lambda\mathcal{P}(X)$  inherited from  $C(X)^*$  with the weak-\* topology. Check that  $\Lambda$  is a homeomorphism when  $ca(X)$  has the narrow topology and  $C(X)^*$  has the weak-\* topology. Then, because the subspace topology  $\Lambda\mathcal{P}(X)$  inherited from  $C(X)^*$  with the weak-\* topology is compact and metrizable, the subspace topology on  $\mathcal{P}(X)$  inherited from  $ca(X)$  with the narrow topology is compact and metrizable.

Suppose that  $\mathcal{P}(X)$  with the narrow topology is compact and metrizable. Because  $\mathcal{P}(X)$  is compact and metrizable it is separable (any compact metrizable topological space is separable), and so  $\{\delta_x : x \in X\}$  with the subspace topology inherited from  $\mathcal{P}(X)$  is separable. But Theorem 6 tells us that there is a homeomorphism between  $X$  and  $\{\delta_x : x \in X\}$ , so  $X$  is separable too. We now know that  $X$  is a separable metrizable space, so by Theorem 6 we get that  $\{\delta_x : x \in X\}$  is a closed subset of  $\mathcal{P}(X)$ , and is therefore compact. Finally, again using that  $X$  and  $\{\delta_x : x \in X\}$  are homeomorphic, we get that  $X$  is compact.  $\square$

The following theorem shows the same type of result as above for separable spaces.<sup>16</sup>

<sup>14</sup>Charalambos D. Aliprantis and Kim C. Border, *Infinite Dimensional Analysis: A Hitchhiker's Guide*, third ed., p. 513, Theorem 15.11.

<sup>15</sup>Charalambos D. Aliprantis and Kim C. Border, *Infinite Dimensional Analysis: A Hitchhiker's Guide*, third ed., p. 239, Theorem 6.30.

<sup>16</sup>Charalambos D. Aliprantis and Kim C. Border, *Infinite Dimensional Analysis: A Hitchhiker's Guide*, third ed., p. 513, Theorem 15.12.



**Theorem 8.** *Suppose that  $X$  is a metrizable topological space. Then  $X$  is separable if and only if  $\mathcal{P}(X)$  with the narrow topology is both separable and metrizable.*

*Proof.* It is a fact that if  $Y$  is a separable metrizable topological space then there is a compatible metric  $d$  on  $Y$  (a metric on a topological space is called **compatible** when it induces the topology) such that  $(Y, d)$  is a **totally bounded metric space**: for every  $\epsilon > 0$ , there are  $y_1, \dots, y_n \in Y$  such that for every  $y \in Y$  there is some  $j$  such that  $d(y, y_j) < \epsilon$ .<sup>17</sup>

Another fact we shall use is the following. If  $(Y, d)$  is a metric space,  $A$  is a nonempty subset of  $Y$ , and  $f : A \rightarrow \mathbb{R}$  is uniformly continuous, then there is a unique uniformly continuous  $\hat{f} : \overline{A} \rightarrow \mathbb{R}$  whose restriction to  $A$  is equal to  $f$ , and this extension satisfies  $\|\hat{f}\|_\infty = \|f\|_\infty$ , and finally  $\widehat{f + g} = \hat{f} + \hat{g}$ .<sup>18</sup>

Suppose that  $X$  is separable, let  $d$  be a compatible metric such that  $(X, d)$  is a totally bounded metric space, and let  $(\hat{X}, \hat{d})$  be the completion of  $(X, d)$ .  $(\hat{X}, \hat{d})$  is compact because a metric space is totally bounded if and only if its completion is compact. Let  $U_d(X)$  denote the set of bounded uniformly continuous functions  $X \rightarrow \mathbb{R}$ . It is apparent that with the supremum norm  $U_d(X)$  is a real normed space, and one proves that it is a closed subset of  $C_b(X)$  and hence itself a Banach space. Define  $\phi : U_d(X) \rightarrow U_{\hat{d}}(\hat{X}) = C(\hat{X})$  by  $\phi(f) = \hat{f}$ , where  $\hat{f}$  is the extension of  $f$  to  $\overline{X} = \hat{X}$  explained in the previous paragraph; the equality  $U_{\hat{d}}(\hat{X}) = C(\hat{X})$  is because  $\hat{X}$  is compact. What we have said so far makes it apparent that  $\phi$  is a linear isometry. Because  $\hat{X}$  is compact, the Banach space  $C(\hat{X})$  is separable and hence the subspace  $\phi(U_d(X))$  is separable. Then, because  $\phi$  is an isometry,  $U_d(X)$  is separable, say with a countable dense subset  $D$ .

It is a fact that if  $Y$  is a set and  $\mathcal{F}$  is a family of functions  $Y \rightarrow \mathbb{R}$  that **separates points in  $Y$**  (if  $x \neq y$  then there is some  $f \in \mathcal{F}$  such that  $f(x) \neq f(y)$ ) then the initial topology on  $Y$  induced by  $\mathcal{F}$  is equal to the subspace topology on  $Y$  inherited from  $\mathbb{R}^{\mathcal{F}}$  with the product topology.<sup>19</sup> One proves that  $D$  being dense in  $U_d(X)$  implies that it separates points in  $\mathcal{P}(X)$ ,<sup>20</sup> and hence that the initial topology on  $\mathcal{P}(X)$  induced by  $D$  is equal to the subspace topology on  $\mathcal{P}(X)$  inherited from  $\mathbb{R}^D$ . But because  $D$  is countable and  $\mathbb{R}$  is separable and metrizable,  $\mathbb{R}^D$  with the product topology is separable and metrizable. As  $\mathbb{R}^D$  is a separable metrizable topological space, the subspace topology on  $\mathcal{P}(X)$  inherited from  $\mathbb{R}^D$  is separable and metrizable. (If  $Y$  is a metrizable topological space and  $A$  is a subset of  $Y$ , then the subspace topology on  $A$  inherited from  $Y$  is metrizable. If  $Y$  is a separable metrizable topological space and  $A$  is a subset of  $Y$ , then the subspace topology on  $A$  inherited from  $Y$

<sup>17</sup>Charalambos D. Aliprantis and Kim C. Border, *Infinite Dimensional Analysis: A Hitchhiker's Guide*, third ed., p. 91, Corollary 3.41.

<sup>18</sup>Charalambos D. Aliprantis and Kim C. Border, *Infinite Dimensional Analysis: A Hitchhiker's Guide*, third ed., p. 77, Lemma 3.11.

<sup>19</sup>Charalambos D. Aliprantis and Kim C. Border, *Infinite Dimensional Analysis: A Hitchhiker's Guide*, third ed., p. 53, Lemma 2.63.

<sup>20</sup>Charalambos D. Aliprantis and Kim C. Border, *Infinite Dimensional Analysis: A Hitchhiker's Guide*, third ed., p. 506, Theorem 15.1.

is separable, but this need not be true if  $Y$  is not metrizable.) This shows that the initial topology on  $\mathcal{P}(X)$  induced by  $D$  is separable and metrizable. But because  $D$  is a dense subset of  $U_d(X)$ , it can be proved that the initial topology on  $\mathcal{P}(X)$  induced by  $D$  is equal to the initial topology on  $\mathcal{P}(X)$  induced by  $C_b(X)$ ,<sup>21</sup> and the initial topology on  $\mathcal{P}(X)$  induced by  $C_b(X)$  is precisely the narrow\* topology on  $\mathcal{P}(X)$ . This shows that the narrow topology on  $\mathcal{P}(X)$  is separable and metrizable.

Suppose that the narrow topology on  $\mathcal{P}(X)$  is separable and metrizable. Then,  $\{\delta_x : x \in X\}$  with the subspace topology inherited from  $\mathcal{P}(X)$  is separable, and by Theorem 6 there is a homeomorphism  $X \rightarrow \{\delta_x : x \in X\}$ , so  $X$  is separable too.  $\square$

The same type of result is true for Polish spaces and Borel spaces: if  $X$  is a metrizable topological space, then  $X$  is a Polish space if and only if  $\mathcal{P}(X)$  with the narrow topology is a Polish space, and  $X$  is a Borel space if and only if  $\mathcal{P}(X)$  with the narrow topology is a Borel space.<sup>22</sup>

Suppose that  $X$  is a Hausdorff space and that  $\mathcal{F}$  is a family of finite positive Borel measures on  $X$ . We say that  $\mathcal{F}$  is **tight** if for every  $\epsilon > 0$  there is a compact subset  $K$  of  $X$  such that each  $\mu \in \mathcal{F}$  satisfies  $\mu(X \setminus K) < \epsilon$ , i.e.  $\mu(K) > \mu(X) - \epsilon$ . We specified that  $X$  be Hausdorff to ensure that any compact subset of  $X$  is a Borel set. Earlier in this note we defined the notion of a tight measure, and saying that a measure  $\mu$  is tight is equivalent to saying that the family  $\{\mu\}$  is tight.

A subset  $A$  of a topological space  $Y$  is said to be **relatively compact** if its closure is a compact set. We prove in the following theorem that every tight family of Borel probability measures on a separable metrizable topological space  $X$  is relatively compact in  $\mathcal{P}(X)$ .<sup>23</sup>

**Theorem 9.** *Suppose that  $X$  is a separable metrizable topological space. Then any tight subset of  $\mathcal{P}(X)$  is relatively compact.*

*Proof.* The **Urysohn metrization theorem** states that for a Hausdorff space  $Y$  the following three statements are equivalent: (i)  $Y$  is separable and metrizable, (ii)  $Y$  is second-countable and regular, (iii)  $Y$  is a subspace of a topological space  $\mathcal{H}$  that is homeomorphic with  $[0, 1]^{\mathbb{N}}$ . Thus,  $X$  is a subspace of a topological space  $\mathcal{H}$  that is homeomorphic with  $[0, 1]^{\mathbb{N}}$ . Except for the last paragraph of the proof, this is the only time where we invoke that  $X$  is separable. In the course of the proof we shall use that  $\mathcal{H}$  is compact and metrizable.

Suppose that  $\mathcal{F}$  is a tight family of Borel probability measures on  $X$ . If  $\mathcal{F} = \emptyset$  then it is immediate that the claim is true. Otherwise, for each  $n \in \mathbb{N}$  let  $\mu_n \in \mathcal{F}$ . For each  $m \in \mathbb{N}$ , because  $\mathcal{F}$  is tight there is a compact subset  $K_m$  of

<sup>21</sup>Charalambos D. Aliprantis and Kim C. Border, *Infinite Dimensional Analysis: A Hitchhiker's Guide*, third ed., p. 507, Theorem 15.2.

<sup>22</sup>Charalambos D. Aliprantis and Kim C. Border, *Infinite Dimensional Analysis: A Hitchhiker's Guide*, third ed., p. 515, Theorem 15.15 and p. 517, Theorem 15.18. A **Borel space** is a topological space that is homeomorphic to a Borel subset of a Polish space.

<sup>23</sup>Charalambos D. Aliprantis and Kim C. Border, *Infinite Dimensional Analysis: A Hitchhiker's Guide*, third ed., p. 518, Lemma 15.21.

$X$  such that  $\mu_n(K_m) > 1 - \frac{1}{m}$  for all  $n \in \mathbb{N}$ . Because  $K_m$  is a compact subset of  $X$  and  $X$  has the subspace topology inherited from  $\mathcal{H}$ ,  $K_m$  is a compact subset of  $\mathcal{H}$  and hence is a Borel set in  $\mathcal{H}$ . (This is worth pointing out, because  $X$  need not be a Borel set in  $\mathcal{H}$ .) Let

$$E = \bigcup_{m \in \mathbb{N}} K_m,$$

which is a Borel set in  $\mathcal{H}$ . We assign  $E$  the subspace topology inherited from  $\mathcal{H}$ . For  $n, m \in \mathbb{N}$ , because  $K_m \subset E$ ,

$$1 \geq \mu_n(E) \geq \mu_n(K_m) > 1 - \frac{1}{m}.$$

This is true for all  $m$ , so  $\mu_n(E) = 1$ , and therefore  $\mu_n \in \mathcal{P}(E)$  for all  $n \in \mathbb{N}$ . For each  $n \in \mathbb{N}$ , define  $\lambda_n : \mathcal{B}_{\mathcal{H}} \rightarrow [0, 1]$  by

$$\lambda_n(B) = \mu_n(E \cap B), \quad B \in \mathcal{B}_{\mathcal{H}}.$$

One checks that  $\lambda_n$  is a positive measure. Because  $\lambda_n(\mathcal{H}) = \mu_n(E) = 1$ , we have  $\lambda_n \in \mathcal{P}(\mathcal{H})$ . The topological space  $\mathcal{H}$  is compact and metrizable, so by Theorem 7, the space  $\mathcal{P}(\mathcal{H})$  with the narrow topology is compact and metrizable, and since  $\lambda_n$  is a sequence in  $\mathcal{P}(\mathcal{H})$ , it has a subsequence  $\lambda_{a(n)}$  that converges to some  $\lambda \in \mathcal{P}(\mathcal{H})$ . For each  $m \in \mathbb{N}$ ,  $K_m$  is a compact subset of  $\mathcal{H}$  and hence closed, so using  $\lambda_{a(n)} \rightarrow \lambda$  in  $\mathcal{P}(\mathcal{H})$  we have by Theorem 5 that

$$\begin{aligned} \lambda(E) &\geq \lambda(K_m) \\ &\geq \limsup_{n \rightarrow \infty} \lambda_{a(n)}(K_m) \\ &= \limsup_{n \rightarrow \infty} \mu_{a(n)}(E \cap K_m) \\ &= \limsup_{n \rightarrow \infty} \mu_{a(n)}(K_m) \\ &\geq \limsup_{n \rightarrow \infty} \left(1 - \frac{1}{m}\right) \\ &= 1 - \frac{1}{m}. \end{aligned}$$

This is true for all  $m$ , so  $\lambda(E) = 1$ , and therefore  $\lambda \in \mathcal{P}(E)$ .

It is a fact that if  $Z$  is a Borel subset of a metrizable topological space  $Y$ , then the narrow topology on  $\mathcal{P}(Z)$  is equal to the subspace topology on  $\mathcal{P}(Z)$  inherited from  $\mathcal{P}(Y)$  with the narrow topology.<sup>24</sup> As  $E$  is a Borel set in  $\mathcal{H}$ , the narrow topology on  $\mathcal{P}(E)$  is equal to the subspace topology on  $\mathcal{P}(E)$  inherited from  $\mathcal{P}(\mathcal{H})$  with the narrow topology. We know that  $\lambda_{a(n)} \rightarrow \lambda$  in  $\mathcal{P}(\mathcal{H})$ , and since the members of the sequence and the limit  $\lambda$  belong to  $\mathcal{P}(E)$ ,  $\lambda_{a(n)} \rightarrow \lambda$

<sup>24</sup>Charalambos D. Aliprantis and Kim C. Border, *Infinite Dimensional Analysis: A Hitchhiker's Guide*, third ed., p. 510, Lemma 15.4.

in  $\mathcal{P}(E)$ . ( $\lambda_n \in \mathcal{P}(E)$  because  $\lambda_n(E) = 1$  for each  $n$ .) But as elements of  $\mathcal{P}(E)$ ,  $\lambda_n = \mu_n$ : for any  $B \in \mathcal{B}_E$ ,  $\lambda_n(B) = \mu_n(E \cap B) = \mu_n(B)$ . So  $\mu_n \rightarrow \lambda$  in  $\mathcal{P}(E)$ . Because  $E$  is a Borel set in  $X$ , the narrow topology on  $\mathcal{P}(E)$  is equal to the subspace topology on  $\mathcal{P}(E)$  inherited from  $\mathcal{P}(X)$  with the narrow topology. We have established that  $\mu_n \rightarrow \lambda$  in  $\mathcal{P}(E)$  with the narrow topology, and therefore  $\mu_n \rightarrow \lambda$  in  $\mathcal{P}(X)$  with the narrow topology. (If  $A$  is a subset of a topological space  $Y$  and a sequence converges in  $A$  with the subspace topology, then the sequence converges in  $Y$  to the same limit.)

Because  $X$  is separable, by Theorem 8 we have that  $\mathcal{P}(X)$  is metrizable (and separable, but we don't care about that here). We have established that any sequence of elements of  $\mathcal{F}$  has a subsequence that converges to some element of  $\mathcal{P}(X)$ , and because  $\mathcal{P}(X)$  this suffices to show that  $\mathcal{F}$  is relatively compact, completing the proof.  $\square$

## 6 The Prokhorov metric on $\mathcal{P}(X)$

Let  $(X, d)$  be a metric space. If  $A \subset X$  and  $x \in X$ , we define

$$d(x, A) = \inf_{a \in A} d(x, a);$$

if  $A = \emptyset$  then  $d(x, A) = \infty$ . For  $A \subset X$  and  $\alpha > 0$ , we define

$$A_\alpha = \{x \in X : d(x, A) < \alpha\}.$$

One checks that for any  $A \subset X$ ,  $\lim_{\epsilon \rightarrow 0} A_\epsilon = \bar{A}$ , the closure of  $A$ .

For  $\mu, \nu \in \mathcal{P}(X)$  we define  $d_P(\mu, \nu)$  to be

$$\inf\{\alpha > 0 : \text{for all } E \in \mathcal{B}_X, \mu(E) \leq \nu(E_\alpha) + \alpha \text{ and } \nu(E) \leq \mu(E_\alpha) + \alpha\}. \quad (2)$$

One proves that  $d_P$  is a metric on  $\mathcal{P}(X)$ , called the **Prokhorov metric**. It is a bounded metric: for any  $E \in \mathcal{B}_X$ ,  $\mu(E), \nu(E) \leq 1$ , so  $\mu(E) \leq \nu(E_1) + 1$  and  $\nu(E) \leq \mu(E_1) + 1$ , hence  $d_P(\mu, \nu) \leq 1$ .

**Theorem 10.** *If  $\mu, \mu_1, \mu_2, \dots \in \mathcal{P}(X)$  and  $d_P(\mu_i, \mu) \rightarrow 0$ , then  $\mu_i$  converges narrowly to  $\mu$ .*

*Proof.* For each  $i$ ,  $d_P(\mu_i, \mu)$  is defined in (2) as an infimum. We inductively define a sequence  $\alpha_i > 0$  by taking  $\alpha_i$  to be (i) an element of the set of which  $d_P(\mu_i, \mu)$  is the infimum, (ii)  $\alpha_{i+1} < \alpha_i$  for each  $i$ , and (iii)  $\alpha_i \rightarrow 0$ ; we can satisfy (iii) because  $d_P(\mu_i, \mu) \rightarrow 0$ . Then for any  $E \in \mathcal{B}_X$ ,

$$\begin{aligned} \limsup_{i \rightarrow \infty} \mu_i(E) &\leq \limsup_{i \rightarrow \infty} \mu(E_{\alpha_i}) + \alpha_i \\ &\leq \limsup_{i \rightarrow \infty} \mu(E_{\alpha_i}) + \limsup_{i \rightarrow \infty} \alpha_i \\ &= \limsup_{i \rightarrow \infty} \mu(E_{\alpha_i}) \\ &= \mu(\bar{A}). \end{aligned}$$

Hence if  $C$  is a closed subset of  $X$ , then, as  $C \in \mathcal{B}_X$ , the above tells us

$$\limsup_{i \rightarrow \infty} \mu_i(C) \leq \mu(C),$$

which shows by Theorem 5 that  $\mu_i$  converges narrowly to  $\mu$ .  $\square$

It follows from the above theorem that the topology on  $\mathcal{P}(X)$  induced by the Prokhorov metric is finer than the narrow topology on  $\mathcal{P}(X)$ .

It can be proved that if  $X$  is a separable metric space then the Prokhorov metric on  $\mathcal{P}(X)$  induces the narrow topology. This is proved in notes by van Gaans.<sup>25</sup> Van Gaans also proves **Prokhorov's theorem** in his notes, which states that if  $X$  is a Polish space, then a subset of  $\mathcal{P}(X)$  is tight if and only if it is relatively compact. We proved one of these implications in Theorem 9 (which is the implication that takes more work to prove) without needing to use that  $X$  is Polish, but only a separable metrizable space. Another result proved in those notes is that if  $X$  is a separable complete metric space, then so is  $\mathcal{P}(X)$  with the Prokhorov metric; it is worth reminding ourselves that this is not immediate from the fact mentioned earlier that if  $X$  is Polish then  $\mathcal{P}(X)$  with the narrow topology is Polish, because even though in this case the narrow topology on  $\mathcal{P}(X)$  is induced by the Prokhorov metric, it need not be the case that a metric that induces completely metrizable topology is itself a complete metric.

## 7 Supports of positive Borel measures

If  $X$  is a topological space and  $\mu$  is a positive Borel measure on  $X$ , the **support of  $\mu$**  is a subset  $S$  of  $X$  such that (i)  $S$  is closed, (ii)  $\mu(X \setminus S) = 0$ , and (iii) if  $U$  is an open set with  $U \cap S \neq \emptyset$  then  $\mu(U \cap S) > 0$ . It is straightforward to check that  $\mu$  has at most one support. We now prove conditions under which a support exists.<sup>26</sup>

**Theorem 11.** *Suppose that  $X$  is a topological space and that  $\mu$  is a positive Borel measure on  $X$ . If  $X$  is second-countable or  $\mu$  is a tight measure, then  $\mu$  has a support.*

*Proof.* Assume that  $(X, \tau)$  is second-countable, and let  $\beta$  be a countable base for  $X$ . Let

$$G = \bigcup_{V \in \beta, \mu(V)=0} V.$$

$G$  is an open set, and because  $\beta$  is countable,

$$\mu(G) \leq \sum_{V \in \beta} \mu(V) = 0.$$

<sup>25</sup>Onno van Gaans, *Probability measures on metric spaces*, <http://www.math.leidenuniv.nl/~vangaans/jancol1.pdf>

<sup>26</sup>Charalambos D. Aliprantis and Kim C. Border, *Infinite Dimensional Analysis: A Hitchhiker's Guide*, third ed., p. 442, Theorem 12.14.

Let  $S = X \setminus G$ , which is closed and has measure 0. Suppose that  $U$  is open, that  $U \cap S \neq \emptyset$ , and that  $\mu(U \cap S) = 0$ . Then

$$\mu(U) = \mu(U \cap S) + \mu(U \cap G) = \mu(U \cap G) \leq \mu(G) = 0.$$

On the other hand, there are  $V_1, V_2, \dots \in \beta$  such that  $U = \bigcap_i V_i$ , and thus for each  $i$ ,  $\mu(V_i) = 0$ , which implies that  $U \subset G$ . But this contradicts that  $U \cap S \neq \emptyset$ , hence  $\mu(U \cap S) > 0$ . Therefore  $S$  is the support of  $\mu$ .

Assume that  $\mu$  is tight and let

$$G = \bigcup_{V \in \tau, \mu(V)=0} V.$$

Then  $G$  is open and  $S = X \setminus G$  is closed. If  $K$  is a compact subset of  $G$ , there are  $V_1, \dots, V_n \in \tau$  with  $\mu(V_i) = 0$  such that  $K \subset \bigcup_{i=1}^n V_i$ , and so  $\mu(K) \leq \sum_{i=1}^n \mu(V_i) = 0$ . Because  $\mu$  is tight,

$$\mu(X \setminus S) = \mu(G) = \sup\{\mu(K) : K \text{ is compact and } K \subset G\},$$

and this supremum is equal to 0. If  $U \in \tau$  and  $U \cap S \neq \emptyset$ , then because  $U$  is not contained in  $G$ ,  $\mu(U) > 0$ . This shows that  $S$  is the support of  $\mu$ .  $\square$