6. Fubini's Theorem

This next example gives a "real world" example of the fact that it is not always possible to interchange order of integration.

Example 6.1. Consider

$$\begin{aligned}\int_0^1 dy \int_1^\infty dx(e^{-xy} - 2e^{-2xy}) &= \int_0^1 dy \left\{ \frac{e^{-y}}{-y} - 2\frac{e^{-2xy}}{-2y} \right\}\Big|_{x=1}^{\infty} \\ &= \int_0^1 dy \left[\frac{e^{-y} - e^{-2y}}{y} \right] \\ &= \int_0^1 dy\ e^{-y} \left(\frac{1-e^{-y}}{y} \right) \in (0,\infty).\end{aligned}$$

Note well that $\left(\frac{1-e^{-y}}{y}\right)$ has not singularity at 0. On the other hand

$$\begin{aligned}\int_1^\infty dx \int_0^1 dy(e^{-xy} - 2e^{-2xy}) &= \int_1^\infty dx \left\{ \frac{e^{-xy}}{-x} - 2\frac{e^{-2xy}}{-2x} \right\}\Big|_{y=0}^{1} \\ &= \int_1^\infty dx \left\{ \frac{e^{-2x} - e^{-x}}{x} \right\} \\ &= -\int_1^\infty e^{-x} \left[\frac{1-e^{-x}}{x} \right] dx \in (-\infty, 0).\end{aligned}$$

Moral $\int dx \int dy\ f(x,y) \neq \int dy \int dx\ f(x,y)$ is **not always true**.

In the remainder of this section we will let $(X,\mathcal{M},\mu)$ and $(Y,\mathcal{N},\nu)$ be fixed measure spaces. Our main goals are to show:

(1) There exists a unique measure $\mu\otimes\nu$ on $\mathcal{M}\otimes\mathcal{N}$ such that $\mu\otimes\nu(A\times B) = \mu(A)\nu(B)$ for all $A\in\mathcal{M}$ and $B\in\mathcal{N}$ and
(2) For all $f : X\times Y \to [0,\infty]$ which are $\mathcal{M}\otimes\mathcal{N}$ – measurable,

$$\begin{aligned}\int_{X\times Y} f\ d(\mu\otimes\nu) &= \int_X d\mu(x) \int_Y d\nu(y) f(x,y) \\ &= \int_Y d\nu(y) \int_X d\mu(x) f(x,y).\end{aligned}$$

Before proving such assertions, we will need a few more technical measure theoretic arguments which are of independent interest.

6.1. Measure Theoretic Arguments.

Definition 6.2. Let $\mathcal{C}\subseteq\mathcal{P}(X)$ be a collection of sets. We say:

(1) $\mathcal{C}$ is a **monotone class** if it is closed under increasing unions and decreasing intersections,
(2) $\mathcal{C}$ is a π – **class** if it is closed under finite intersections and
(3) $\mathcal{C}$ is a λ–**class** if $\mathcal{C}$ satisfies the following properties:
   (a) $X\in\mathcal{C}$
   (b) If $A,B\in\mathcal{C}$ and $A\cap B=\emptyset$, then $A\cup B\in\mathcal{C}$. (Closed under disjoint unions.)
   (c) If $A,B\in\mathcal{C}$ and $A\supset B$, then $A\setminus B\in\mathcal{C}$. (Closed under proper differences.)

(d) If $A_n \in \mathcal{C}$ and $A_n \uparrow A$, then $A \in \mathcal{C}$. (Closed under countable increasing unions.)

(4) We will say $\mathcal{C}$ is a λ_0 – class if $\mathcal{C}$ satisfies conditions a) – c) but not necessarily d).

Remark 6.3. Notice that every λ – class is also a monotone class. a λ – class is a monotone class.

Lemma 6.4 (Monotone Class Theorem)**.** *Suppose $\mathcal{A} \subseteq \mathcal{P}(X)$ is an algebra and $\mathcal{C}$ is the smallest monotone class containing $\mathcal{A}$. Then $\mathcal{C} = \sigma(\mathcal{A})$.*

Proof. For $C \in \mathcal{C}$ let

$$\mathcal{C}(C) = \{B \in \mathcal{C} : C \cap B, C \cap B^c, B \cap C^c \in \mathcal{C}\},$$

then $\mathcal{C}(C)$ is a monotone class. Indeed, if $B_n \in \mathcal{C}(C)$ and $B_n \uparrow B$, then $B_n^c \downarrow B^c$ and so

$$\begin{aligned} &\mathcal{C} \ni C \cap B_n \uparrow C \cap B \\ &\mathcal{C} \ni C \cap B_n^c \downarrow C \cap B^c \text{ and} \\ &\mathcal{C} \ni B_n \cap C^c \uparrow B \cap C^c. \end{aligned}$$

Since $\mathcal{C}$ is a monotone class, it follows that $C \cap B, C \cap B^c, B \cap C^c \in \mathcal{C}$, i.e. $B \in \mathcal{C}(C)$. This shows that $\mathcal{C}(C)$ is closed under increasing limits and a similar argument shows that $\mathcal{C}(C)$ is closed under decreasing limits. Thus we have shown that $\mathcal{C}(C)$ is a monotone class for all $C \in \mathcal{C}$.

If $A \in \mathcal{A} \subset \mathcal{C}$, then $A \cap B, A \cap B^c, B \cap A^c \in \mathcal{A} \subset \mathcal{C}$ for all $B \in \mathcal{A}$ and hence it follows that $\mathcal{A} \subset \mathcal{C}(A) \subset \mathcal{C}$. Since $\mathcal{C}$ is the smallest monotone class containing $\mathcal{A}$ and $\mathcal{C}(A)$ is a monotone class containing $\mathcal{A}$, we conclude that $\mathcal{C}(A) = \mathcal{C}$ for any $A \in \mathcal{A}$.

Let $B \in \mathcal{C}$ and notice that $A \in \mathcal{C}(B)$ happens iff $B \in \mathcal{C}(A)$. This observation and the fact that $\mathcal{C}(A) = \mathcal{C}$ for all $A \in \mathcal{A}$ implies $\mathcal{A} \subset \mathcal{C}(B) \subset \mathcal{C}$ for all $B \in \mathcal{C}$. Again since $\mathcal{C}$ is the smallest monotone class containing $\mathcal{A}$ and $\mathcal{C}(B)$ is a monotone class we conclude that $\mathcal{C}(B) = \mathcal{C}$ for all $B \in \mathcal{C}$. That is to say, if $A, B \in \mathcal{C}$ then $A \in \mathcal{C} = \mathcal{C}(B)$ and hence $A \cap B$, $A \cap B^c$, $A^c \cap B \in \mathcal{C}$. So $\mathcal{C}$ is closed under complements (since $X \in \mathcal{A} \subset \mathcal{C}$) and finite intersections and increasing unions from which it easily follows that $\mathcal{C}$ is a σ – algebra. ■

Let $\mathcal{E} \subset \mathcal{P}(X \times Y)$ be given by

$$\mathcal{E} = \mathcal{M} \times \mathcal{N} = \{A \times B : A \in \mathcal{M}, B \in \mathcal{N}\}$$

and recall from Exercise 4.2 that $\mathcal{E}$ is an elementary family. Hence the algebra $\mathcal{A} = \mathcal{A}(\mathcal{E})$ generated by $\mathcal{E}$ consists of sets which may be written as disjoint unions of sets from $\mathcal{E}$.

Theorem 6.5 (Uniqueness)**.** *Suppose that $\mathcal{E} \subset \mathcal{P}(X)$ is an elementary class and $\mathcal{M} = \sigma(\mathcal{E})$ (the σ – algebra generated by $\mathcal{E}$). If μ and ν are two measures on $\mathcal{M}$ which are σ – finite on $\mathcal{E}$ and such that $\mu = \nu$ on $\mathcal{E}$ then $\mu = \nu$ on $\mathcal{M}$.*

Proof. Let $\mathcal{A} := \mathcal{A}(\mathcal{E})$ be the algebra generated by $\mathcal{E}$. Since every element of $\mathcal{A}$ is a disjoint union of elements from $\mathcal{E}$, it is clear that $\mu = \nu$ on $\mathcal{A}$. Henceforth we may assume that $\mathcal{E} = \mathcal{A}$. We begin first with the special case where $\mu(X) < \infty$ and hence $\nu(X) = \mu(X) < \infty$. Let

$$\mathcal{C} = \{A \in \mathcal{M} : \mu(A) = \nu(A)\}$$

The reader may easily check that $\mathcal{C}$ is a monotone class. Since $\mathcal{A} \subset \mathcal{C}$, the monotone class lemma asserts that $\mathcal{M} = \sigma(\mathcal{A}) \subset \mathcal{C} \subset \mathcal{M}$ showing that $\mathcal{C} = \mathcal{M}$ and hence that $\mu = \nu$ on $\mathcal{M}$.

For the σ – finite case, let $X_n \in \mathcal{A}$ be sets such that $\mu(X_n) = \nu(X_n) < \infty$ and $X_n \uparrow X$ as $n \to \infty$. For $n \in \mathbb{N}$, let

$$\mu_n(A) := \mu(A \cap X_n) \text{ and } \nu_n(A) = \nu(A \cap X_n) \tag{6.1}$$

for all $A \in \mathcal{M}$. Then one easily checks that μ_n and ν_n are finite measure on $\mathcal{M}$ such that $\mu_n = \nu_n$ on $\mathcal{A}$. Therefore, by what we have just proved, $\mu_n = \nu_n$ on $\mathcal{M}$. Hence or all $A \in \mathcal{M}$, using the continuity of measures,

$$\mu(A) = \lim_{n\to\infty} \mu(A \cap X_n) = \lim_{n\to\infty} \nu(A \cap X_n) = \nu(A).$$

■

Lemma 6.6. *If $\mathcal{D}$ is a λ_0 – class which contains a π-class, $\mathcal{C}$, then $\mathcal{D}$ contains $\mathcal{A}(\mathcal{C})$ – the algebra generated by $\mathcal{C}$.*

Proof. We will give two proofs of this lemma. The first proof is "constructive" and makes use of Proposition 4.9 which tells how to construct $\mathcal{A}(\mathcal{C})$ from $\mathcal{C}$. The key to the first proof is the following claim which will be proved by induction.

Claim. Let $\tilde{\mathcal{C}}_0 = \mathcal{C}$ and $\tilde{\mathcal{C}}_n$ denote the collection of subsets of X of the form

$$A_1^c \cap \cdots \cap A_n^c \cap B = B \setminus A_1 \setminus A_2 \setminus \cdots \setminus A_n. \tag{6.2}$$

with $A_i \in \mathcal{C}$ and $B \in \mathcal{C} \cup \{X\}$. Then $\tilde{\mathcal{C}}_n \subset \mathcal{D}$ for all n, i.e. $\tilde{\mathcal{C}} := \cup_{n=0}^{\infty} \tilde{\mathcal{C}}_n \subset \mathcal{D}$.

By assumption $\tilde{\mathcal{C}}_0 \subset \mathcal{D}$ and when $n = 1$,

$$B \setminus A_1 = B \setminus (A_1 \cap B) \in \mathcal{D}$$

when $A_1, B \in \mathcal{C} \subset \mathcal{D}$ since $A_1 \cap B \in \mathcal{C} \subset \mathcal{D}$. Therefore, $\tilde{\mathcal{C}}_1 \subset \mathcal{D}$. For the induction step, let $B \in \mathcal{C} \cup \{X\}$ and $A_i \in \mathcal{C} \cup \{X\}$ and let E_n denote the set in Eq. (6.2) We now assume $\tilde{\mathcal{C}}_n \subset \mathcal{D}$ and wish to show $E_{n+1} \in \mathcal{D}$, where

$$E_{n+1} = E_n \setminus A_{n+1} = E_n \setminus (A_{n+1} \cap E_n).$$

Because

$$A_{n+1} \cap E_n = A_1^c \cap \cdots \cap A_n^c \cap (B \cap A_{n+1}) \in \tilde{\mathcal{C}}_n \subset \mathcal{D}$$

and $(A_{n+1} \cap E_n) \subset E_n \in \tilde{\mathcal{C}}_n \subset \mathcal{D}$, we have $E_{n+1} \in \mathcal{D}$ as well. This finishes the proof of the claim.

Notice that $\tilde{\mathcal{C}}$ is still a multiplicative class and from Proposition 4.9 (using the fact that $\mathcal{C}$ is a multiplicative class), $\mathcal{A}(\mathcal{C})$ consists of finite unions of elements from $\tilde{\mathcal{C}}$. By applying the claim to $\tilde{\mathcal{C}}$, $A_1^c \cap \cdots \cap A_n^c \in \mathcal{D}$ for all $A_i \in \tilde{\mathcal{C}}$ and hence

$$A_1 \cup \cdots \cup A_n = (A_1^c \cap \cdots \cap A_n^c)^c \in \mathcal{D}.$$

Thus we have shown $\mathcal{A}(\mathcal{C}) \subset \mathcal{D}$ which completes the proof.

(**Second Proof.**) With out loss of generality, we may assume that $\mathcal{D}$ is the smallest λ_0 – class containing $\mathcal{C}$ for if not just replace $\mathcal{D}$ by the intersection of all λ_0 – classes containing $\mathcal{C}$. Let

$$\mathcal{D}_1 := \{A \in \mathcal{D} : A \cap C \in \mathcal{D}\ \forall\ C \in \mathcal{C}\}.$$

Then $\mathcal{C} \subset \mathcal{D}_1$ and $\mathcal{D}_1$ is also a λ_0–class as we now check. a) $X \in \mathcal{D}_1$. b) If $A, B \in \mathcal{D}_1$ with $A \cap B = \emptyset$, then $(A \cup B) \cap C = (A \cap C) \coprod (B \cap C) \in \mathcal{D}$ for all $C \in \mathcal{C}$. c) If $A, B \in \mathcal{D}_1$ with $B \subset A$, then $(A \setminus B) \cap C = A \cap C \setminus (B \cap C) \in \mathcal{D}$ for all $C \in \mathcal{C}$. Since

$\mathcal{C} \subset \mathcal{D}_1 \subset \mathcal{D}$ and $\mathcal{D}$ is the smallest λ_0 – class containing $\mathcal{C}$ it follows that $\mathcal{D}_1 = \mathcal{D}$. From this we conclude that if $A \in \mathcal{D}$ and $B \in \mathcal{C}$ then $A \cap B \in \mathcal{D}$.

Let

$$\mathcal{D}_2 := \{A \in \mathcal{D} : A \cap D \in \mathcal{D} \; \forall \; D \in \mathcal{D}\}.$$

Then $\mathcal{D}_2$ is a λ_0–class (as you should check) which, by the above paragraph, contains $\mathcal{C}$. As above this implies that $\mathcal{D} = \mathcal{D}_2$, i.e. we have shown that $\mathcal{D}$ is closed under finite intersections. Since λ_0 – classes are closed under complementation, $\mathcal{D}$ is an algebra and hence $\mathcal{A}(\mathcal{C}) \subset \mathcal{D}$. In fact $\mathcal{D} = \mathcal{A}(\mathcal{C})$. ■

This Lemma along with the monotone class theorem immediately implies Dynkin's very useful "$\pi - \lambda$ theorem."

Theorem 6.7 ($\pi - \lambda$ Theorem). *If $\mathcal{D}$ is a λ class which contains a contains a π-class, $\mathcal{C}$, then $\sigma(\mathcal{C}) \subset \mathcal{D}$.*

Proof. Since $\mathcal{D}$ is a λ_0 – class, Lemma 6.6 implies that $\mathcal{A}(\mathcal{C}) \subset \mathcal{D}$ and so by Remark 6.3 and Lemma 6.4, $\sigma(\mathcal{C}) \subset \mathcal{D}$. Let us pause to give a second stand-alone proof of this Theorem.

(**Second Proof**.) With out loss of generality, we may assume that $\mathcal{D}$ is the smallest λ – class containing $\mathcal{C}$ for if not just replace $\mathcal{D}$ by the intersection of all λ – classes containing $\mathcal{C}$. Let

$$\mathcal{D}_1 := \{A \in \mathcal{D} : A \cap C \in \mathcal{D} \; \forall \; C \in \mathcal{C}\}.$$

Then $\mathcal{C} \subset \mathcal{D}_1$ and $\mathcal{D}_1$ is also a λ–class because as we now check. a) $X \in \mathcal{D}_1$. b) If $A, B \in \mathcal{D}_1$ with $A \cap B = \emptyset$, then $(A \cup B) \cap C = (A \cap C) \coprod (B \cap C) \in \mathcal{D}$ for all $C \in \mathcal{C}$. c) If $A, B \in \mathcal{D}_1$ with $B \subset A$, then $(A \setminus B) \cap C = A \cap C \setminus (B \cap C) \in \mathcal{D}$ for all $C \in \mathcal{C}$. d) If $A_n \in \mathcal{D}_1$ and $A_n \uparrow A$ as $n \to \infty$, then $A_n \cap C \in \mathcal{D}$ for all $C \in \mathcal{D}$ and hence $A_n \cap C \uparrow A \cap C \in \mathcal{D}$. Since $\mathcal{C} \subset \mathcal{D}_1 \subset \mathcal{D}$ and $\mathcal{D}$ is the smallest λ – class containing $\mathcal{C}$ it follows that $\mathcal{D}_1 = \mathcal{D}$. From this we conclude that if $A \in \mathcal{D}$ and $B \in \mathcal{C}$ then $A \cap B \in \mathcal{D}$.

Let

$$\mathcal{D}_2 := \{A \in \mathcal{D} : A \cap D \in \mathcal{D} \; \forall \; D \in \mathcal{D}\}.$$

Then $\mathcal{D}_2$ is a λ–class (as you should check) which, by the above paragraph, contains $\mathcal{C}$. As above this implies that $\mathcal{D} = \mathcal{D}_2$, i.e. we have shown that $\mathcal{D}$ is closed under finite intersections.

Since λ – classes are closed under complementation, $\mathcal{D}$ is an algebra which is closed under increasing unions and hence is closed under arbitrary countable unions, i.e. $\mathcal{D}$ is a σ – algebra. Since $\mathcal{C} \subset \mathcal{D}$ we must have $\sigma(\mathcal{C}) \subset \mathcal{D}$ and in fact $\sigma(\mathcal{C}) = \mathcal{D}$. ■

Using this theorem we may strengthen Theorem 6.8 to the following.

Theorem 6.8 (Uniqueness). *Suppose that $\mathcal{C} \subset \mathcal{P}(X)$ is a π – class such that $\mathcal{M} = \sigma(\mathcal{C})$. If μ and ν are two measures on $\mathcal{M}$ and there exists $X_n \in \mathcal{C}$ such that $X_n \uparrow X$ and $\mu(X_n) = \nu(X_n) < \infty$ for each n, then $\mu = \nu$ on $\mathcal{M}$.*

Proof. As in the proof of Theorem 6.5, it suffices to consider the case where μ and ν are finite measure such that $\mu(X) = \nu(X) < \infty$. In this case the reader may easily verify from the basic properties of measures that

$$\mathcal{D} = \{A \in \mathcal{M} : \mu(A) = \nu(A)\}$$

is a λ – class. By assumption $\mathcal{C} \subset \mathcal{D}$ and hence by the π– λ theorem, $\mathcal{D}$ contains $\mathcal{M} = \sigma(\mathcal{C})$. ■

As an immediate consequence we have the following corollaries.

Corollary 6.9. *Suppose that (X, τ) is a topological space, $\mathcal{M} = \sigma(\tau)$ is the Borel σ – algebra on X and μ and ν are two measures on $\mathcal{M}$ which are σ – finite on τ. If $\mu = \nu$ on τ then $\mu = \nu$ on $\mathcal{M}$, i.e. $\mu \equiv \nu$.*

Corollary 6.10. *Suppose that μ and ν are two measures on $\mathcal{B}_{\mathbb{R}^n}$ which are finite on bounded sets and such that $\mu(A) = \nu(A)$ for all sets A of the form*

$$A = (a, b] = (a_1, b_1] \times \cdots \times (a_n, b_n]$$

with $a, b \in \mathbb{R}^n$ and $a \leq b$, i.e. $a_i \leq b_i$ for all i. Then $\mu = \nu$ on $\mathcal{B}_{\mathbb{R}^n}$.

To end this section we wish to reformulate the $\pi - \lambda$ theorem in a function theoretic setting.

Theorem 6.11. *Let X be a set and $\mathcal{H}$ be a subspace of $B(X, \mathbb{R})$ – the space of bounded real valued functions on X. Assume:*

1. *$1 \in \mathcal{H}$, i.e. the constant functions are in $\mathcal{H}$ and*
2. *$\mathcal{H}$ is closed under bounded convergence, i.e. if $\{f_n\}_{n=1}^{\infty} \subset \mathcal{H}$ is a sequence of functions such that $M := \sup_n \|f_n\|_\infty < \infty$ and $f(x) := \lim_{n\to\infty} f_n(x)$ exists for all $x \in X$ then $f \in \mathcal{H}$.*

If $\mathcal{C} \subset \mathcal{P}(X)$ is a multiplicative class such that $1_A \in \mathcal{H}$ for all $A \in \mathcal{C}$, then $\mathcal{H}$ contains all bounded $\sigma(\mathcal{C})$ – measurable functions.

Proof. Let $\mathcal{D} := \{A \subset X : 1_A \in \mathcal{H}\}$. Then by assumption $\mathcal{C} \subset \mathcal{D}$ and since $1 \in \mathcal{H}$ we know $X \in \mathcal{D}$. If $A, B \in \mathcal{D}$ are disjoint then $1_{A\cup B} = 1_A + 1_B \in \mathcal{H}$ so that $A \cup B \in \mathcal{D}$ and if $A, B \in \mathcal{D}$ and $A \subset B$, then $1_{B\setminus A} = 1_B - 1_A \in \mathcal{H}$. Finally if $A_n \in \mathcal{D}$ and $A_n \uparrow A$ as $n \to \infty$ then $1_{A_n} \to 1_A$ boundedly so $1_A \in \mathcal{H}$ and hence $A \in \mathcal{D}$. So $\mathcal{D}$ is λ – class containing $\mathcal{C}$ and hence $\mathcal{D}$ contains $\sigma(\mathcal{C})$. From this it follows that $\mathcal{H}$ contains 1_A for all $A \in \sigma(\mathcal{C})$ and hence all $\sigma(\mathcal{C})$ – measurable simple functions by linearity. The proof is now complete with an application of the approximation Theorem 5.13 along with the assumption that $\mathcal{H}$ is closed under bounded convergence. ■

Corollary 6.12. *Suppose that (X, d) is a metric space and $\mathcal{M} = \sigma(\tau_d)$ is the Borel σ – algebra on X and $\mathcal{H}$ is a subspace of $B(X, \mathbb{R})$ such that $C_b(X, \mathbb{R}) \subset \mathcal{H}$ ($C_b(X, \mathbb{R})$ – the bounded continuous functions on X) and $\mathcal{H}$ is closed under bounded convergence. Then $\mathcal{H}$ contains all bounded $\mathcal{M}$ – measurable real valued functions on X. (This may be paraphrased as follows. The smallest vector space of bounded functions which is closed under bounded convergence and contains $C_b(X, \mathbb{R})$ is the space of bounded $\mathcal{M}$ – measurable real valued functions on X.)*

Proof. Let $V \in \tau_d$ be an open subset of X and for $n \in \mathbb{N}$ let

$$f_n(x) := \min(n \cdot d_{V^c}(x), 1) \text{ for all } x \in X.$$

Notice that $f_n = \phi_n \circ d_{V^c}$ where $\phi_n(t) = \min(nt, 1)$ which is continuous and hence $f_n \in C_b(X, \mathbb{R})$ for all n. Furthermore, f_n converges boundedly to 1_V as $n \to \infty$ and therefore $1_V \in \mathcal{H}$ for all $V \in \tau$. Since τ is a π – class the corollary follows by an application of Theorem 6.11. ■

Here is a basic application of this corollary.

Proposition 6.13. *Suppose that (X,d) is a metric space, μ and ν are two measures on $\mathcal{M}=\sigma(\tau_d)$ which are finite on bounded measurable subsets of X and*

$$\int_X f d\mu = \int_X f d\nu \tag{6.3}$$

for all $f \in C_b(X,\mathbb{R})$ such that $\operatorname{supp}(f)$ *is bounded. Then $\mu \equiv \nu$.*

Proof. To prove this fix a $o \in X$ and let

$$\psi_R(x) = ([R+1-d(x,o)] \wedge 1) \vee 0$$

so that $\psi_R \in C_b(X,[0,1])$, $\operatorname{supp}(\psi_R) \subset B(o,R+2)$ and $\psi_R \uparrow 1$ as $R \to \infty$. Let $\mathcal{H}_R$ denote the space of bounded measurable functions f such that

$$\int_X \psi_R f d\mu = \int_X \psi_R f d\nu. \tag{6.4}$$

Then $\mathcal{H}$ is closed under bounded convergence and because of Eq. (6.3) contains $C_b(X,\mathbb{R})$. Therefore by Corollary 6.12, $\mathcal{H}_R$ contains all bounded measurable functions on X. Take $f = 1_A$ in Eq. (6.4) with $A \in \mathcal{M}$, and then use the monotone convergence theorem to let $R \to \infty$. The result is $\mu(A) = \nu(A)$ for all $A \in \mathcal{M}$. ■

6.2. Fubini-Tonelli's Theorem and Product Measure.

Recall that $(X,\mathcal{M},\mu)$ and $(Y,\mathcal{N},\nu)$ are fixed measure spaces.

Notation 6.14. Suppose that $f : X \to \mathbb{C}$ and $g : Y \to \mathbb{C}$ are functions, let $f \otimes g$ denote the function on $X \times Y$ given by

$$f \otimes g(x,y) = f(x)g(y).$$

Notice that if f, g are measurable, then $f \otimes g$ is $(\mathcal{M} \otimes \mathcal{N}, \mathcal{B}_{\mathbb{C}})$ – measurable. To prove this let $F(x,y) = f(x)$ and $G(x,y) = g(y)$ so that $f \otimes g = F \cdot G$ will be measurable provided that F and G are measurable. Now $F = f \circ \pi_1$ where $\pi_1 : X \times Y \to X$ is the projection map. This shows that F is the composition of measurable functions and hence measurable. Similarly one shows that G is measurable.

Theorem 6.15. *Suppose $(X,\mathcal{M},\mu)$ and $(Y,\mathcal{N},\nu)$ are σ-finite measure spaces and f is a nonnegative $(\mathcal{M} \otimes \mathcal{N}, \mathcal{B}_{\mathbb{R}})$ – measurable function, then for each $y \in Y$,*

$$x \to f(x,y) \text{ is } \mathcal{M} - \mathcal{B}_{[0,\infty]} \text{ measurable,} \tag{6.5}$$

for each $x \in X$,

$$y \to f(x,y) \text{ is } \mathcal{N} - \mathcal{B}_{[0,\infty]} \text{ measurable,} \tag{6.6}$$

$$x \to \int_Y f(x,y) d\nu(y) \text{ is } \mathcal{M} - \mathcal{B}_{[0,\infty]} \text{ measurable,} \tag{6.7}$$

$$y \to \int_X f(x,y) d\mu(x) \text{ is } \mathcal{N} - \mathcal{B}_{[0,\infty]} \text{ measurable,} \tag{6.8}$$

and

$$\int_X d\mu(x) \int_Y d\nu(y) f(x,y) = \int_Y d\nu(y) \int_X d\mu(x) f(x,y). \tag{6.9}$$

Proof. Suppose that $E = A \times B \in \mathcal{E} := \mathcal{M} \times \mathcal{N}$ and $f = 1_E$. Then

$$f(x,y) = 1_{A\times B}(x,y) = 1_A(x)1_B(y)$$

and one sees that Eqs. (6.5) and (6.6) hold. Moreover

$$\int_Y f(x,y)d\nu(y) = \int_Y 1_A(x)1_B(y)d\nu(y) = 1_A(x)\nu(B),$$

so that Eq. (6.7) holds and we have

$$\int_X d\mu(x) \int_Y d\nu(y) f(x,y) = \nu(B)\mu(A).$$

Similarly,

$$\int_X f(x,y)d\mu(x) = \mu(A)1_B(y) \text{ and}$$
$$\int_Y d\nu(y) \int_X d\mu(x) f(x,y) = \nu(B)\mu(A)$$

from which it follows that Eqs. (6.8) and (6.9) hold in this case as well.

For the moment let us further assume that $\mu(X) < \infty$ and $\nu(Y) < \infty$ and let H be the collection of all bounded $(\mathcal{M}\otimes\mathcal{N}, \mathcal{B}_{\mathbb{R}})$ – measurable functions on $X\times Y$ such that Eqs. (6.5) – (6.9) hold. Using the fact that measurable functions are closed under pointwise limits and the dominated convergence theorem (the dominating function always being a constant), one easily shows that H closed under bounded convergence. Since we have just verified that $1_E \in H$ for all E in the π – class, $\mathcal{E}$, it follows that H is the space of all bounded $(\mathcal{M}\otimes\mathcal{N}, \mathcal{B}_{\mathbb{R}})$ – measurable functions on $X \times Y$. Finally if $f : X \times Y \to [0,\infty]$ is a $(\mathcal{M}\otimes\mathcal{N}, \mathcal{B}_{\bar{\mathbb{R}}})$ – measurable function, let $f_M = M \wedge f$ so that $f_M \uparrow f$ as $M \to \infty$ and Eqs. (6.5) – (6.9) hold with f replaced by f_M for all $M \in \mathbb{N}$. Repeated use of the monotone convergence theorem allows us to pass to the limit $M \to \infty$ in these equations to deduce the theorem in the case μ and ν are finite measures.

For the σ – finite case, choose $X_n \in \mathcal{M}$, $Y_n \in \mathcal{N}$ such that $X_n \uparrow X$, $Y_n \uparrow Y$, $\mu(X_n) < \infty$ and $\nu(Y_n) < \infty$ for all $m, n \in \mathbb{N}$. Then define $\mu_m(A) = \mu(X_m \cap A)$ and $\nu_n(B) = \nu(Y_n \cap B)$ for all $A \in \mathcal{M}$ and $B \in \mathcal{N}$ or equivalently $d\mu_m = 1_{X_m} d\mu$ and $d\nu_n = 1_{Y_n} d\nu$. By what we have just proved Eqs. (6.5) – (6.9) with μ replaced by μ_m and ν by ν_n for all $(\mathcal{M}\otimes\mathcal{N}, \mathcal{B}_{\bar{\mathbb{R}}})$ – measurable functions, $f : X \times Y \to [0,\infty]$. The validity of Eqs. (6.5) – (6.9) then follows by passing to the limits $m \to \infty$ and then $n \to \infty$ using the monotone convergence theorem again to conclude

$$\int_X f d\mu_m = \int_X f 1_{X_m} d\mu \uparrow \int_X f d\mu \text{ as } m \to \infty$$

and

$$\int_Y g d\mu_n = \int_Y g 1_{Y_n} d\mu \uparrow \int_Y g d\mu \text{ as } n \to \infty$$

for all $f \in L^+(X,\mathcal{M})$ and $g \in L^+(Y,\mathcal{N})$. ■

Corollary 6.16. *Suppose $(X,\mathcal{M},\mu)$ and $(Y,\mathcal{N},\nu)$ are σ-finite measure spaces. Then there exists a unique measure π on $\mathcal{M}\otimes\mathcal{N}$ such that $\pi(A\times B) = \mu(A)\nu(B)$ for all $A \in \mathcal{M}$ and $B \in \mathcal{N}$. Moreover π is given by*

$$\pi(E) = \int_X d\mu(x) \int_Y d\nu(y) 1_E(x,y) = \int_Y d\nu(y) \int_X d\mu(x) 1_E(x,y) \tag{6.10}$$

for all $E \in \mathcal{M} \otimes \mathcal{N}$ and π is σ – finite.

Notation 6.17. The measure π is called the product measure of μ and ν and will be denoted by $\mu \otimes \nu$.

Proof. Notice that any measure π such that $\pi(A \times B) = \mu(A)\nu(B)$ for all $A \in \mathcal{M}$ and $B \in \mathcal{N}$ is necessarily σ – finite. Indeed, let $X_n \in \mathcal{M}$ and $Y_n \in \mathcal{N}$ be chosen so that $\mu(X_n) < \infty$, $\nu(Y_n) < \infty$, $X_n \uparrow X$ and $Y_n \uparrow Y$, then $X_n \times Y_n \in \mathcal{M} \otimes \mathcal{N}$, $X_n \times Y_n \uparrow X \times Y$ and $\pi(X_n \times Y_n) < \infty$ for all n. The uniqueness assertion is a consequence of either Theorem 6.5 or by Theorem 6.8 with $\mathcal{E} = \mathcal{M} \times \mathcal{N}$. For the existence, it suffices to observe, using the monotone convergence theorem, that π defined in Eq. (6.10) is a measure on $\mathcal{M} \otimes \mathcal{N}$. Moreover this measure satisfies $\pi(A \times B) = \mu(A)\nu(B)$ for all $A \in \mathcal{M}$ and $B \in \mathcal{N}$ from Eq. (55.10). ∎

Theorem 6.18 (Tonelli's Theorem)**.** *Suppose $(X, \mathcal{M}, \mu)$ and $(Y, \mathcal{N}, \nu)$ are σ-finite measure spaces and $\pi = \mu \otimes \nu$ be the product measure on $\mathcal{M} \otimes \mathcal{N}$. If $f \in L^+(X \times Y, \mathcal{M} \otimes \mathcal{N})$, then $f(\cdot, y) \in L^+(X, \mathcal{M})$ for all $y \in Y$, $f(x, \cdot) \in L^+(Y, \mathcal{N})$ for all $x \in X$,*

$$\int_Y f(\cdot, y) d\nu(y) \in L^+(X, \mathcal{M}), \int_X f(x, \cdot) d\mu(x) \in L^+(Y, \mathcal{N})$$

and

$$\int_{X \times Y} f \, d\pi = \int_X d\mu(x) \int_Y d\nu(y) f(x, y) \tag{6.11}$$

$$= \int_Y d\nu(y) \int_X d\mu(x) f(x, y). \tag{6.12}$$

Proof. By Theorem 6.15 and Corollary 6.16, the theorem holds when $f = 1_E$ with $E \in \mathcal{M} \otimes \mathcal{N}$. Using the linearity of all of the statements, the theorem is also true for non-negative simple functions. Then using the monotone convergence theorem repeatedly along with Theorem 5.13, one deduces the theorem for general $f \in L^+(X \times Y, \mathcal{M} \otimes \mathcal{N})$. ∎

Theorem 6.19 (Fubini's Theorem)**.** *Suppose $(X, \mathcal{M}, \mu)$ and $(Y, \mathcal{N}, \nu)$ are σ-finite measure spaces and $\pi = \mu \otimes \nu$ be the product measure on $\mathcal{M} \otimes \mathcal{N}$. If $f \in L^1(\pi)$ then for μ a.e. x, $f(x, \cdot) \in L^1(\nu)$ and for ν a.e. y, $f(\cdot, y) \in L^1(\mu)$. Moreover,*

$$g(x) = \int_Y f(x, y) dv(y) \text{ and } h(y) = \int_X f(x, y) d\mu(x)$$

are in $L^1(\mu)$ and $L^1(\nu)$ respectively and Eq. (6.12) holds.

Proof. If $f \in L^1(X \times Y) \cap L^+$ then by Eq. (6.11),

$$\int_X \left(\int_Y f(x, y) d\nu(y) \right) d\mu(x) < \infty$$

so $\int_Y f(x, y) d\nu(y) < \infty$ for μ a.e. x, i.e. for μ a.e. x, $f(x, \cdot) \in L^1(\nu)$. Similarly for ν a.e. y, $f(\cdot, y) \in L^1(\mu)$. Let f be a real valued function in $f \in L^1(X \times Y)$ and let $f = f_+ - f_-$. Apply the results just proved to $f_\pm$ to conclude, $f_\pm(x, \cdot) \in L^1(\nu)$ for μ a.e. x and that

$$\int_Y f_\pm(\cdot, y) d\nu(y) \in L^1(\mu).$$

Therefore for μ a.e. .x,

$$f(x, \cdot) = f_+(x, \cdot) - f_-(x, \cdot) \in L^1(\nu)$$

and

$$x \to \int f(x,y)d\nu(y) = \int f_+(x,\cdot)d\nu(y) - \int f_-(x,\cdot)d\nu(y)$$

is a μ – almost everywhere defined function such that $\int f(\cdot,y)d\nu(y) \in L^1(\mu)$. Because

$$\int f_\pm(x,y)d\left(\mu\otimes\nu\right) = \int d\mu(x)\int d\nu(y)f_\pm(x,y),$$

$$\begin{aligned}\int f\,d(\mu\otimes\nu) &= \int f_+ d(\mu\otimes\nu) - \int f_- d(\mu\otimes\nu)\\ &= \int d\mu\int d\nu f_+ - \int d\mu\int d\nu f_-\\ &= \int d\mu\left(\int f_+ d\nu - \int f_- d\nu\right)\\ &= \int d\mu\int d\nu(f_+ - f_-) = \int d\mu\int d\nu f.\end{aligned}$$

The proof that

$$\int f\,d(\mu\otimes\nu) = \int d\nu(y)\int d\mu(x)f(x,y)$$

is analogous. As usual the complex case follows by applying the real results just proved to the real and imaginary parts of f. ■

Notation 6.20. Given $E \subset X\times Y$ and $x \in X$, let

$${}_xE := \{y\in Y : (x,y)\in E\}.$$

Similarly if $y \in Y$ is given let

$$E_y := \{x\in X : (x,y)\in E\}.$$

If $f : X\times Y \to \mathbb{C}$ is a function let $f_x = f(x,\cdot)$ and $f^y := f(\cdot,y)$ so that $f_x : Y \to \mathbb{C}$ and $f^y : X \to \mathbb{C}$.

Theorem 6.21. *Suppose $(X,\mathcal{M},\mu)$ and $(Y,\mathcal{N},\nu)$ are complete σ-finite measure spaces. Let $(X\times Y,\mathcal{L},\lambda)$ be the completion of $(X\times Y,\mathcal{M}\otimes\mathcal{N},\mu\otimes\nu)$. If f is $\mathcal{L}$-measurable and (a) $f\geq 0$ or (b) $f\in L^1(\lambda)$ then f_x is $\mathcal{N}$-measurable for μ a.e. x and f^y is $\mathcal{M}$-measurable for ν a.e. y and in case (b) $f_x\in L^1(\nu)$ and $f^y \in L^1(\mu)$ for μ a.e. x and ν a.e. y respectively. Moreover,*

$$x \to \int f_x d\nu \text{ and } y \to \int f^y d\mu$$

are measurable and

$$\int f d\lambda = \int d\nu\int d\mu\, f = \int d\mu\int d\nu\, f.$$

Proof. If $E \in \mathcal{M}\otimes\mathcal{N}$ is a $\mu\otimes\nu$ null set ($(\mu\otimes\nu)(E) = 0$), then

$$0 = (\mu\otimes\nu)(E) = \int_X \nu({}_xE)d\mu(x) = \int_X \mu(E_y)d\nu(y).$$

This shows that

$$\mu(\{x : \nu({}_xE)\neq 0\}) = 0 \text{ and } \nu(\{y : \mu(E_y)\neq 0\}) = 0,$$

i.e. $\nu({}_xE) = 0$ for μ a.e. x and $\mu(E_y) = 0$ for ν a.e. y.

If h is $\mathcal{L}$ measurable and $h = 0$ for λ- a.e., then there exists $E \in \mathcal{M} \otimes \mathcal{N} \ni \{(x, y) : h(x, y) \neq 0\} \subseteq E$ and $(\mu \otimes \nu)(E) = 0$. Therefore $|h(x, y)| \leq 1_E(x, y)$ and $(\mu \otimes \nu)(E) = 0$. Since

$$\{h_x \neq 0\} = \{y \in Y : h(x, y) \neq 0\} \subset {}_xE \text{ and}$$
$$\{h_y \neq 0\} = \{x \in X : h(x, y) \neq 0\} \subset E_y$$

we learn that for μ a.e. x and ν a.e. y that $\{h_x \neq 0\} \in \mathcal{M}$, $\{h_y \neq 0\} \in \mathcal{N}$, $\nu(\{h_x \neq 0\}) = 0$ and a.e. and $\mu(\{h_y \neq 0\}) = 0$. This implies

$$\text{for } \nu \text{ a.e. } y, \int h(x, y) d\nu(y) \text{ exists and equals } 0$$
$$\text{and}$$
$$\text{for } \mu \text{ a.e. } x, \int h(x, y) d\mu(y) \text{ exists and equals } 0.$$

Therefore

$$0 = \int h d\lambda = \int \left(\int h d\mu \right) d\nu = \int \left(\int h d\nu \right) d\mu.$$

For general $f \in L^1(\lambda)$, we may choose $g \in L^1(\mathcal{M} \otimes \mathcal{N}, \mu \otimes \nu)$ such that $f(x, y) = g(x, y)$ for $\lambda-$ a.e. (x, y). Define $h \equiv f - g$. Then $h = 0$, $\lambda-$ a.e. Hence by what we have just proved and Theorem 6.18 $f = g + h$ has the following properties:

(1) For μ a.e. x, $y \to f(x, y) = g(x, y) + h(x, y)$ is in $L^1(\nu)$ and
$$\int f(x, y) d\nu(y) = \int g(x, y) d\nu(y).$$
(2) For ν a.e. y, $x \to f(x, y) = g(x, y) + h(x, y)$ is in $L^1(\mu)$ and
$$\int f(x, y) d\mu(x) = \int g(x, y) d\mu(x).$$

From these assertions and Theorem 6.18, it follows that

$$\begin{aligned}
\int d\mu(x) \int d\nu(y) f(x, y) &= \int d\mu(x) \int d\nu(y) g(x, y) \\
&= \int d\nu(y) \int d\nu(x) g(x, y) \\
&= \int g(x, y) d(\mu \otimes \nu)(x, y) \\
&= \int f(x, y) d\lambda(x, y)
\end{aligned}$$

and similarly we shows

$$\int d\nu(y) \int d\mu(x) f(x, y) = \int f(x, y) d\lambda(x, y).$$

■

The previous theorems have obvious generalizations to products of any finite number of σ – compact measure spaces. For example the following theorem holds.

Theorem 6.22. *Suppose $\{(X_i, \mathcal{M}_i, \mu_i)\}_{i=1}^n$ are σ – finite measure spaces and $X := X_1 \times \cdots \times X_n$. Then there exists a unique measure, π, on $(X, \mathcal{M}_1 \otimes \cdots \otimes \mathcal{M}_n)$ such that $\pi(A_1 \times \cdots \times A_n) = \mu_1(A_1) \ldots \mu_n(A_n)$ for all $A_i \in \mathcal{M}_i$. (This measure and its* *completion will be denote by $\mu_1 \otimes \cdots \otimes \mu_n$.) If $f : X \to [0,\infty]$ is a measurable function then*

$$\int_X f d\pi = \prod_{i=1}^n \int_{X_{\sigma(i)}} d\mu_{\sigma(i)}(x_{\sigma(i)})\ f(x_1,\dots,x_n)$$

where σ is any permutation of $\{1,2,\dots,n\}$. This equation also holds for any $f \in L^1(X,\pi)$ and moreover, $f \in L^1(X,\pi)$ iff

$$\prod_{i=1}^n \int_{X_{\sigma(i)}} d\mu_{\sigma(i)}(x_{\sigma(i)})\ |f(x_1,\dots,x_n)| < \infty$$

for some (and hence all) permutation, σ.

This theorem can be proved by the same methods as in the two factor case. Alternatively, one can use induction on n, see Exercise 6.6.

Example 6.23. For any $\Lambda, M \geq 0$,

$$\int_0^\infty \frac{\sin x}{x} e^{-\Lambda x} dx = \frac{1}{2}\pi - \arctan \Lambda \text{ for all } \Lambda > 0 \tag{6.13}$$

and

$$\left| \int_0^M \frac{\sin x}{x} e^{-\Lambda x} dx - \frac{1}{2}\pi + \arctan \Lambda \right| \leq C \frac{e^{-M\Lambda}}{M} \tag{6.14}$$

where $C = \max_{x\geq 0} \frac{1+x}{1+x^2} = \frac{1}{2\sqrt{2}-2} \cong 1.2$. In particular,

$$\lim_{M\to\infty} \int_0^M \frac{\sin x}{x} dx = \pi/2. \tag{6.15}$$

To verify these assertions, first notice that by the fundamental theorem of calculus,

$$|\sin x| = \left| \int_0^x \cos y dy \right| \leq \left| \int_0^x |\cos y|\, dy \right| \leq \left| \int_0^x 1 dy \right| = |x|$$

so $\left|\frac{\sin x}{x}\right| \leq 1$ for all $x \neq 0$. Making use of the identity

$$\int_0^\infty e^{-tx} dt = 1/x$$

and Fubini's theorem,

$$\begin{aligned}
\int_0^M \frac{\sin x}{x} e^{-\Lambda x} dx &= \int_0^M dx \sin x\, e^{-\Lambda x} \int_0^\infty e^{-tx} dt \\
&= \int_0^\infty dt \int_0^M dx \sin x\, e^{-(\Lambda+t)x} \\
&= \int_0^\infty \frac{1 - (\cos M + (\Lambda+t)\sin M)\, e^{-M(\Lambda+t)}}{(\Lambda+t)^2+1} dt \\
&= \int_0^\infty \frac{1}{(\Lambda+t)^2+1} dt - \int_0^\infty \frac{\cos M + (\Lambda+t)\sin M}{(\Lambda+t)^2+1} e^{-M(\Lambda+t)} dt \\
&= \frac{1}{2}\pi - \arctan \Lambda - \epsilon(M,\Lambda) \qquad (6.16)
\end{aligned}$$

where

$$\epsilon(M,\Lambda)=\int_0^\infty \frac{\cos M+(\Lambda+t)\sin M}{(\Lambda+t)^2+1}e^{-M(\Lambda+t)}dt.$$

Since

$$\left|\frac{\cos M+(\Lambda+t)\sin M}{(\Lambda+t)^2+1}\right|\le\frac{1+(\Lambda+t)}{(\Lambda+t)^2+1}\le C,$$

$$|\epsilon(M,\Lambda)|\le\int_0^\infty e^{-M(\Lambda+t)}dt=C\frac{e^{-M\Lambda}}{M}.$$

This estimate along with Eq. (6.16) proves Eq. (6.14) from which Eq. (6.15) follows by taking $\Lambda\to\infty$ and Eq. (6.13) follows (using the dominated convergence theorem again) by letting $M\to\infty$.

6.3. Lebesgue measure on $\mathbb{R}^d$.

Notation 6.24. Let

$$m^d:=\overbrace{m\otimes\cdots\otimes m}^{d\text{ times}}\text{ on }\mathcal{B}_{\mathbb{R}^d}=\overbrace{\mathcal{B}_{\mathbb{R}}\otimes\cdots\otimes\mathcal{B}_{\mathbb{R}}}^{d\text{ times}}$$

be the d – fold product of Lebesgue measure m on $\mathcal{B}_{\mathbb{R}}$. We will also use m^d to denote its completion and let $\mathcal{L}_d$ be the completion of $\mathcal{B}_{\mathbb{R}^d}$ relative to m. A subset $A\in\mathcal{L}_d$ is called a Lebesgue measurable set and m^d is called d – dimensional Lebesgue measure, or just Lebesgue measure for short.

Definition 6.25. A function $f:\mathbb{R}^d\to\mathbb{R}$ is **Lebesgue measurable** if $f^{-1}(\mathcal{B}_{\mathbb{R}})\subseteq\mathcal{L}_d$.

Theorem 6.26. *Lebesgue measure m^d is translation invariant. Moreover m^d is the unique translation invariant measure on $\mathcal{B}_{\mathbb{R}^d}$ such that $m^d((0,1]^d)=1$.*

Proof. Let $A=J_1\times\cdots\times J_d$ with $J_i\in\mathcal{B}_{\mathbb{R}}$ and $x\in\mathbb{R}^d$. Then

$$x+A=(x_1+J_1)\times(x_2+J_2)\times\cdots\times(x_d+J_d)$$

and therefore by translation invariance of m on $\mathcal{B}_{\mathbb{R}}$ we find that

$$m^d(x+A)=m(x_1+J_1)\dots m(x_d+J_d)=m(J_1)\dots m(J_d)=m^d(A)$$

and hence $m^d(x+A)=m^d(A)$ for all $A\in\mathcal{B}_{\mathbb{R}^d}$ by Corollary 6.10. From this fact we see that the measure $m^d(x+\cdot)$ and $m^d(\cdot)$ have the same null sets. Using this it is easily seen that $m(x+A)=m(A)$ for all $A\in\mathcal{L}_d$. The proof of the second assertion is Exercise 6.7. ■

Notation 6.27. I will often be sloppy in the sequel and write m for m^d and dx for $dm(x)=dm^d(x)$. Hopefully the reader will understand the meaning from the context.

The following change of variable theorem is an important tool in using Lebesgue measure.

Theorem 6.28 (Change of Variables Theorem)**.** *Let $\Omega\subset_o\mathbb{R}^d$ be an open set and $T:\Omega\to T(\Omega)\subset_o\mathbb{R}^d$ be a C^1 – diffeomorphism*[11]*. Then for any Borel measurable*

[11] That is $T:\Omega\to T(\Omega)\subset_o\mathbb{R}^d$ is a continuously differentiable bijection and the inverse map $T^{-1}:T(\Omega)\to\Omega$ is also continuously differentiable.

$f : T(\Omega) \to [0, \infty]$,

$$\int_\Omega f \circ T |\det T'| dm = \int_{T(\Omega)} f \; dm, \tag{6.17}$$

where $T'(x)$ is the linear transformation on $\mathbb{R}^d$ defined by $T'(x)v := \frac{d}{dt}|_0 T(x + tv)$. Alternatively, the ij – matrix entry of $T'(x)$ is given by $T'(x)_{ij} = \partial T_j(x)/\partial x_i$ where $T(x) = (T_1(x), \ldots, T_d(x))$.

We will postpone the full proof of this theorem until Section 12. However we will give here the proof in the case that T is linear. The following elementary remark will be used in the proof.

Remark 6.29. Suppose that

$$\Omega \xrightarrow{T} T(\Omega) \xrightarrow{S} S(T(\Omega))$$

are two C^1 – diffeomorphisms and Theorem 6.28 holds for T and S separately, then it holds for the composition $S \circ T$. Indeed

$$\begin{aligned}\int_\Omega f \circ S \circ T |\det (S \circ T)'| dm &= \int_\Omega f \circ S \circ T |\det (S' \circ T) T'| dm \\ &= \int_\Omega (|\det S'| \, f \circ S) \circ T |\det T'| dm \\ &= \int_{T(\Omega)} |\det S'| \, f \circ S dm = \int_{S(T(\Omega))} f dm.\end{aligned}$$

Theorem 6.30. *Suppose $T \in GL(d, \mathbb{R})$ – the space of $d \times d$ invertible matrices.*

(1) *If $f : \mathbb{R}^d \to \mathbb{R}$ is Borel – measurable then so is $f \circ T$ and if $f \geq 0$ or $f \in L^1$ then*

$$\int f(y) dy = |\det T| \int f \circ T(x) dx. \tag{6.18}$$

(2) *If $E \in \mathcal{L}_d$ then $T(E) \in \mathcal{L}_d$ and $m(T(E)) = |\det T| \, m(E)$.*

Proof. Since f is Borel measurable and $T : \mathbb{R}^d \to \mathbb{R}^d$ is continuous and hence Borel measurable, $f \circ T$ is also Borel measurable. We now break the proof of Eq. (6.18) into a number of cases. In each case we make use Tonelli's theorem and the basic properties of one dimensional Lebesgue measure.

(1) Suppose that $i < k$ and

$$T(x_1, x_2 \ldots, x_d) = (x_1, \ldots, x_{i-1}, x_k, x_{i+1} \ldots, x_{k-1}, x_i, x_{k+1}, \ldots x_d)$$

then by Tonelli's theorem,

$$\begin{aligned}\int f \circ T(x_1, \ldots, x_d) &= \int f(x_1, \ldots, x_k, \ldots x_i, \ldots x_d) dx_1 \ldots dx_d \\ &= \int f(x_1, \ldots, x_d) dx_1 \ldots dx_d\end{aligned}$$

which prove Eq. (6.18) in this case since $|\det T| = 1$.

(2) Suppose that $c \in \mathbb{R}$ and $T(x_1, \ldots x_k, \ldots, x_d) = (x_1, \ldots, cx_k, \ldots x_d)$, then

$$\begin{aligned}\int f \circ T(x_1, \ldots, x_d) dm &= \int f(x_1, \ldots, cx_k, \ldots x_i, \ldots x_d) dx_1 \ldots dx_k \ldots dx_d \\ &= |c|^{-1} \int f(x_1, \ldots, x_d) dx_1 \ldots dx_d \\ &= |\det T|^{-1} \int f \; dm\end{aligned}$$

which again proves Eq. (6.18) in this case.

(3) Suppose that

$$T(x_1, x_2 \ldots, x_d) = (x_1, \ldots, \overset{\text{i'th spot}}{x_i + cx_k}, \ldots x_k, \ldots x_d).$$

Then

$$\begin{aligned}\int f \circ T(x_1, \ldots, x_d) dm &= \int f(x_1, \ldots, x_i + cx_k, \ldots x_k, \ldots x_d) dx_1 \ldots dx_i \ldots dx_k \ldots dx_d \\ &= \int f(x_1, \ldots, x_i, \ldots x_k, \ldots x_d) dx_1 \ldots dx_i \ldots dx_k \ldots dx_d \\ &= \int f(x_1, \ldots, x_d) dx_1 \ldots dx_d\end{aligned}$$

where in the second inequality we did the x_i integral first and used translation invariance of Lebesgue measure. Again this proves Eq. (6.18) in this case since $\det(T) = 1$.

Since every invertible matrix is a product of matrices of the type occurring in steps 1. – 3. above, it follows by Remark 6.29 that Eq. (6.18) holds in general. For the second assertion, let $E \in \mathcal{B}_{\mathbb{R}^d}$ and take $f = 1_E$ in Eq. (6.18) to find

$$|\det T| m(T^{-1}(E)) = |\det T| \int 1_{T^{-1}(E)} dm = |\det T| \int 1_E \circ T dm = \int 1_E dm = m(E).$$

Replacing T by T^{-1} in this equation shows that

$$m(T(E)) = |\det T| \, m(E)$$

for all $E \in \mathcal{B}_{\mathbb{R}^d}$. In particular this shows that $m \circ T$ and m have the same null sets and therefore the completion of $\mathcal{B}_{\mathbb{R}^d}$ is $\mathcal{L}_d$ for both measures. Using Proposition 5.7 one now easily shows

$$m(T(E)) = |\det T| \, m(E) \; \forall \; E \in \mathcal{L}_d.$$

■

6.4. **Polar Coordinates and Surface Measure.** Let

$$S^{d-1} = \{x \in \mathbb{R}^d : |x|^2 := \sum_{i=1}^d x_i^2 = 1\}$$

be the unit sphere in $\mathbb{R}^d$. Let $\Phi : \mathbb{R}^d \setminus (0) \to (0, \infty) \times S^{d-1}$ and Φ^{-1} be the inverse map given by

$$\Phi(x) := (|x|, \frac{x}{|x|}) \text{ and } \Phi^{-1}(r, \omega) = r\omega \tag{6.19}$$

respectively. Since Φ and Φ^{-1} are continuous, they are Borel measurable.

Consider the measure $\Phi_* m$ on $\mathcal{B}_{(0,\infty)} \otimes \mathcal{B}_{S^{d-1}}$ given by

$$\Phi_* m(A) := m\left(\Phi^{-1}(A)\right)$$

for all $A \in \mathcal{B}_{(0,\infty)} \otimes \mathcal{B}_{S^{d-1}}$. For $E \in \mathcal{B}_{S^{d-1}}$ and $a > 0$, let

$$E_a := \{r\omega : r \in (0, a] \text{ and } \omega \in E\} = \Phi^{-1}((0, a] \times E) \in \mathcal{B}_{\mathbb{R}^d}.$$

Noting that $E_a = aE_1$, we have for $0 < a < b$, $E \in \mathcal{B}_{S^{d-1}}, E$ and $A = (a, b] \times E$ that

$$\Phi^{-1}(A) = \{r\omega : r \in (a, b] \text{ and } \omega \in E\} \tag{6.20}$$

$$= bE_1 \setminus aE_1. \tag{6.21}$$

Therefore,

$$\begin{aligned} (\Phi_* m)\left((a, b] \times E\right) &= m\left(bE_1 \setminus aE_1\right) = m(bE_1) - m(aE_1) \\ &= b^d m(E_1) - a^d m(E_1) \\ &= d \cdot m(E_1) \int_a^b r^{d-1} dr. \end{aligned} \tag{6.22}$$

Let ρ denote the unique measure on $\mathcal{B}_{(0,\infty)}$ such that

$$\rho(J) = \int_J r^{d-1} dr \tag{6.23}$$

for all $J \in \mathcal{B}_{(0,\infty)}$. Symbolically, we will abbreviate this by writing $\rho(dr) = r^{d-1} dr$.

Definition 6.31. For $E \in \mathcal{B}_{S^{d-1}}$, let $\sigma(E) := d \cdot m(E_1)$. We call σ the surface measure on S.

It is easy to check that σ is a measure. Indeed if $E \in \mathcal{B}_{S^{d-1}}$, then $E_1 = \Phi^{-1}\left((0, 1] \times E\right) \in \mathcal{B}_{\mathbb{R}^d}$ so that $m(E_1)$ is well defined. Moreover if $E = \coprod_{i=1}^\infty E_i$, then $E_1 = \coprod_{i=1}^\infty (E_i)_1$ and

$$\sigma(E) = d \cdot m(E_1) = \sum_{i=1}^\infty m\left((E_i)_1\right) = \sum_{i=1}^\infty \sigma(E_i).$$

The intuition behind this definition is as follows. If $E \subset S^{d-1}$ is a set and $\epsilon > 0$ is a small number, then the volume of

$$(1, 1 + \epsilon] \cdot E = \{r\omega : r \in (1, 1 + \epsilon] \text{ and } \omega \in E\}$$

should be approximately given by $m\left((1, 1 + \epsilon] \cdot E\right) \cong \sigma(E)\epsilon$. On the other hand

$$m\left((1, 1 + \epsilon]E\right) = m\left(E_{1+\epsilon} \setminus E_1\right) = \left\{(1 + \epsilon)^d - 1\right\} m(E_1).$$

Therefore we expect the area of E should be given by

$$\sigma(E) = \lim_{\epsilon \downarrow 0} \frac{\left\{(1 + \epsilon)^d - 1\right\} m(E_1)}{\epsilon} = d \cdot m(E_1).$$

According to these definitions and Eq. (6.22) we have shown that

$$\Phi_* m((a, b] \times E) = \rho((a, b]) \cdot \sigma(E). \tag{6.24}$$

Let

$$\mathcal{E} = \left\{(a, b] \times E : 0 < a < b, E \in \mathcal{B}_{S^{d-1}}\right\},$$

then $\mathcal{E}$ is an elementary class. Since $\sigma(\mathcal{E}) = \mathcal{B}_{(0,\infty)} \otimes \mathcal{B}_{S^{d-1}}$, we conclude from Eq. (6.24) that

$$\Phi_* m = \rho \otimes \sigma$$

and this implies the following theorem.

Theorem 6.32. *If $f : \mathbb{R}^d \to [0,\infty]$ is a $(\mathcal{B}_{R^d}, \mathcal{B})$–measurable function then*

$$\int f(x)dm(x) = \int\limits_{[0,\infty)\times S} f(r\ \omega)\ d\sigma(\omega) r^{d-1} dr. \tag{6.25}$$

Let us now work out some integrals using Eq. (6.25).

Lemma 6.33. *Let $a > 0$ and*

$$I_d(a) := \int\limits_{\mathbb{R}^d} e^{-a|x|^2} dm(x).$$

Then $I_d(a) = (\pi/a)^{d/2}$.

Proof. By Tonelli's theorem and induction,

$$\begin{aligned} I_d(a) &= \int_{\mathbb{R}^{d-1}\times\mathbb{R}} e^{-a|y|^2} e^{-at^2} m_{d-1}(dy)\ dt \\ &= I_{d-1}(a) I_1(a) = I_1^d(a). \end{aligned} \tag{6.26}$$

So it suffices to compute:

$$I_2(a) = \int\limits_{\mathbb{R}^2} e^{-a|x|^2} dm(x) = \int\limits_{\mathbb{R}^2\setminus\{0\}} e^{-a(x_1^2+x_2^2)} dx_1 dx_2.$$

We now make the change of variables, .

$$x_1 = r\cos\theta \text{ and } x_2 = r\sin\theta \text{ for } 0 < r < \infty \text{ and } 0 < \theta < 2\pi.$$

In vector form this transform is

$$x = T(r,\theta) = \begin{pmatrix} r\cos\theta \\ r\sin\theta \end{pmatrix}$$

and the differential and the Jacobian determinant are given by

$$T'(r,\theta) = \begin{pmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{pmatrix} \text{ and } \det T'(r,\theta) = r\cos^2\theta + r\sin^2\theta = r.$$

Notice that $T : (0,\infty) \times (0, 2\pi) \to \mathbb{R}^2 \setminus \ell$ where ℓ is the ray, $\ell := \{(x,0) : x \geq 0\}$ which is a m^2 – null set. Hence by Tonelli's theorem and the change of variable theorem, for any Borel measurable function $f : \mathbb{R}^2 \to [0,\infty]$ we have

$$\int_{\mathbb{R}^2} f(x)dx = \int_0^{2\pi} \int_0^\infty f(r\cos\theta, r\sin\theta)\, r dr d\theta.$$

In particular,

$$\begin{aligned} I_2(a) &= \int_0^\infty dr\ r \int_0^{2\pi} d\theta\ e^{-ar^2} = 2\pi \int_0^\infty r e^{-ar^2} dr \\ &= 2\pi \lim_{M\to\infty} \int_0^M r e^{-ar^2} dr = 2\pi \lim_{M\to\infty} \frac{e^{-ar^2}}{-2a} \int_0^M = \frac{2\pi}{2a} = \pi/a. \end{aligned}$$

This shows that $I_2(a) = \pi/a$ and the result now follows from Eq. (6.26). ■

Corollary 6.34. *The surface area* $\sigma(S^{d-1})$ *of the unit sphere* $S^{d-1}\subset\mathbb{R}^d$ *is*

$$\sigma(S^{d-1})=\frac{2\pi^{d/2}}{\Gamma(d/2)} \tag{6.27}$$

where Γ *is the gamma function given by*

$$\Gamma(x):=\int_0^\infty r^{x-1}e^{-r}dr \tag{6.28}$$

Moreover, $\Gamma(1/2)=\sqrt{\pi}$ *and* $\Gamma(1)=1$.

Proof. We may alternatively compute $I_d(1)=\pi^{d/2}$ using Theorem 6.32;

$$\begin{aligned} I_d(1) &= \int_0^\infty dr\ r^{d-1}e^{-r^2}\int_{S^{d-1}} d\sigma \\ &= \sigma(S^{d-1})\int_0^\infty r^{d-1}e^{-r^2}dr. \end{aligned}$$

We simplify this last integral by making the change of variables $u=r^2$ so that $r=u^{1/2}$ and $dr=\frac{1}{2}u^{-1/2}du$. The result is

$$\begin{aligned} \int_0^\infty r^{d-1}e^{-r^2}dr &= \int_0^\infty u^{\frac{d-1}{2}}e^{-u}\frac{1}{2}u^{-1/2}du \\ &= \frac{1}{2}\int_0^\infty u^{\frac{d}{2}-1}e^{-u}du \\ &= \frac{1}{2}\Gamma(d/2). \end{aligned} \tag{6.29}$$

Collecting these observations implies that

$$\pi^{d/2}=I_d(1)=\frac{1}{2}\sigma(S^{d-1})\Gamma(d/2)$$

which proves Eq. (6.27).

The computation of $\Gamma(1)$ is easy and is left to the reader. By Eq. (6.29),

$$\begin{aligned} \Gamma(1/2) &= 2\int_0^\infty e^{-r^2}dr=\int_{-\infty}^\infty e^{-r^2}dr \\ &= I_1(1)=\sqrt{\pi}. \end{aligned}$$

■

6.5. **Regularity of Measures Theorems.**

Definition 6.35. Suppose that $\mathcal{E}$ is a collection of subsets of X, let $\mathcal{E}_\sigma$ denote the collection of subsets of X which are finite or countable unions of sets from $\mathcal{E}$. Similarly let $\mathcal{E}_\delta$ denote the collection of subsets of X which are finite or countable intersections of sets from $\mathcal{E}$. We also write $\mathcal{E}_{\sigma\delta}=(\mathcal{E}_\sigma)_\delta$ and $\mathcal{E}_{\delta\sigma}=(\mathcal{E}_\delta)_\sigma$, etc.

Remark 6.36. Notice that if $C=\cup C_i$ and $D=\cup D_j$ with $C_i,D_j\in\mathcal{A}_\sigma$, then

$$C\cap D=\cup_{i,j}(C_i\cap D_j)\in\mathcal{A}_\sigma$$

so that $\mathcal{A}_\sigma$ is closed under finite intersections.

The following theorem shows how recover a measure μ on $\sigma(\mathcal{A})$ from its values on an algebra $\mathcal{A}$.

Theorem 6.37 (Regularity Theorem)**.** *Let $\mathcal{A} \subseteq \mathcal{P}(X)$ be an algebra of sets, $\mathcal{M} = \sigma(\mathcal{A})$ and $\mu : \mathcal{M} \to [0,\infty]$ be a measure on $\mathcal{M}$ which is σ – finite on $\mathcal{A}$. Then for all $A \in \mathcal{M}$,*

$$\mu(A) = \inf \{\mu(B) : A \subseteq B \in \mathcal{A}_\sigma\}. \tag{6.30}$$

Moreover, if $A \in \mathcal{M}$ and $\epsilon > 0$ are given, then there exists $B \in \mathcal{A}_\sigma$ such that $A \subset B$ and $\mu(B \setminus A) \leq \epsilon$.

Proof. For $A \subset X$, define

$$\mu^*(A) = \inf \{\mu(B) : A \subseteq B \in \mathcal{A}_\sigma\}.$$

We are trying to show $\mu^* = \mu$ on $\mathcal{M}$. We will begin by first assuming that μ is a finite measure, i.e. $\mu(X) < \infty$.

Let

$$\mathcal{F} = \{B \in \mathcal{M} : \mu^*(B) = \mu(B)\} = \{B \in \mathcal{M} : \mu^*(B) \leq \mu(B)\}.$$

It is clear that $\mathcal{A} \subset \mathcal{F}$, so the finite case will be finished by showing $\mathcal{F}$ is a monotone class. Suppose $B_n \in \mathcal{F}$, $B_n \uparrow B$ as $n \to \infty$ and let $\epsilon > 0$ be given. Since $\mu^*(B_n) = \mu(B_n)$ there exists $A_n \in \mathcal{A}_\sigma$ such that $B_n \subset A_n$ and $\mu(A_n) \leq \mu(B_n) + \epsilon 2^{-n}$ i.e.

$$\mu(A_n \setminus B_n) \leq \epsilon 2^{-n}.$$

Let $A = \cup_n A_n \in \mathcal{A}_\sigma$, then $B \subset A$ and

$$\begin{aligned}\mu(A \setminus B) = \mu(\cup_n (A_n \setminus B)) &\leq \sum_{n=1}^\infty \mu((A_n \setminus B)) \\ &\leq \sum_{n=1}^\infty \mu((A_n \setminus B_n)) \leq \sum_{n=1}^\infty \epsilon 2^{-n} = \epsilon.\end{aligned}$$

Therefore,

$$\mu^*(B) \leq \mu(A) \leq \mu(B) + \epsilon$$

and since $\epsilon > 0$ was arbitrary it follows that $B \in \mathcal{F}$.

Now suppose that $B_n \in \mathcal{F}$ and $B_n \downarrow B$ as $n \to \infty$ so that

$$\mu(B_n) \downarrow \mu(B) \text{ as } n \to \infty.$$

As above choose $A_n \in \mathcal{A}_\sigma$ such that $B_n \subset A_n$ and

$$0 \leq \mu(A_n) - \mu(B_n) = \mu(A_n \setminus B_n) \leq 2^{-n}.$$

Combining the previous two equations shows that $\lim_{n\to\infty} \mu(A_n) = \mu(B)$. Since $\mu^*(B) \leq \mu(A_n)$ for all n, we conclude that $\mu^*(B) \leq \mu(B)$, i.e. that $B \in \mathcal{F}$.

Since $\mathcal{F}$ is a monotone class containing the algebra $\mathcal{A}$, the monotone class theorem asserts that

$$\mathcal{M} = \sigma(\mathcal{A}) \subset \mathcal{F} \subset \mathcal{M}$$

showing the $\mathcal{F} = \mathcal{M}$ and hence that $\mu^* = \mu$ on $\mathcal{M}$.

For the σ – finite case, let $X_n \in \mathcal{A}$ be sets such that $\mu(X_n) < \infty$ and $X_n \uparrow X$ as $n \to \infty$. Let μ_n be the finite measure on $\mathcal{M}$ defined by $\mu_n(A) := \mu(A \cap X_n)$ for all $A \in \mathcal{M}$. Suppose that $\epsilon > 0$ and $A \in \mathcal{M}$ are given. By what we have just proved, for all $A \in \mathcal{M}$, there exists $B_n \in \mathcal{A}_\sigma$ such that $A \subset B_n$ and

$$\mu((B_n \cap X_n) \setminus (A \cap X_n)) = \mu_n(B_n \setminus A) \leq \epsilon 2^{-n}.$$

Notice that since $X_n \in \mathcal{A}_\sigma$, $B_n \cap X_n \in \mathcal{A}_\sigma$ and

$$B := \cup_{n=1}^\infty (B_n \cap X_n) \in \mathcal{A}_\sigma.$$

Moreover, $A \subset B$ and

$$\mu(B \setminus A) \leq \sum_{n=1}^{\infty} \mu((B_n \cap X_n) \setminus A) \leq \sum_{n=1}^{\infty} \mu((B_n \cap X_n) \setminus (A \cap X_n))$$
$$\leq \sum_{n=1}^{\infty} \epsilon 2^{-n} = \epsilon.$$

Since this implies that

$$\mu(A) \leq \mu(B) \leq \mu(A) + \epsilon$$

and $\epsilon > 0$ is arbitrary, this equation shows that Eq. (6.30) holds. ■

Corollary 6.38. *Let $\mathcal{A} \subseteq \mathcal{P}(X)$ be an algebra of sets, $\mathcal{M} = \sigma(\mathcal{A})$ and $\mu : \mathcal{M} \to [0, \infty]$ be a measure on $\mathcal{M}$ which is σ – finite on $\mathcal{A}$. Then for all $A \in \mathcal{M}$ and $\epsilon > 0$ there exists $B \in \mathcal{A}_\delta$ such that $B \subset A$ and*

$$\mu(A \setminus B) < \epsilon.$$

Furthermore, for any $B \in \mathcal{M}$ there exists $A \in \mathcal{A}_{\delta\sigma}$ and $C \in \mathcal{A}_{\sigma\delta}$ such that $A \subset B \subset C$ and $\mu(C \setminus A) = 0$.

Proof. By Theorem 6.37, there exist $C \in \mathcal{A}_\sigma$ such that $A^c \subset C$ and $\mu(C \setminus A^c) \leq \epsilon$. Let $B = C^c \subset A$ and notice that $B \in \mathcal{A}_\delta$ and that $C \setminus A^c = B^c \cap A = A \setminus B$, so that

$$\mu(A \setminus B) = \mu(C \setminus A^c) \leq \epsilon.$$

Finally, given $B \in \mathcal{M}$, we may choose $A_n \in \mathcal{A}_\delta$ and $C_n \in \mathcal{A}_\sigma$ such that $A_n \subset B \subset C_n$ and $\mu(C_n \setminus B) \leq 1/n$ and $\mu(B \setminus A_n) \leq 1/n$. By replacing A_N by $\cup_{n=1}^N A_n$ and C_N by $\cap_{n=1}^N C_n$, we may assume that $A_n \uparrow$ and $C_n \downarrow$ as n increases. Let $A = \cup A_n \in \mathcal{A}_{\delta\sigma}$ and $C = \cap C_n \in \mathcal{A}_{\sigma\delta}$, then $A \subset B \subset C$ and

$$\mu(C \setminus A) = \mu(C \setminus B) + \mu(B \setminus A) \leq \mu(C_n \setminus B) + \mu(B \setminus A_n)$$
$$\leq 2/n \to 0 \text{ as } n \to \infty.$$

■

Corollary 6.39. *Let $\mathcal{A} \subseteq \mathcal{P}(X)$ be an algebra of sets, $\mathcal{M} = \sigma(\mathcal{A})$ and $\mu : \mathcal{M} \to [0, \infty]$ be a measure on $\mathcal{M}$ which is σ – finite on $\mathcal{A}$. Then for every $B \in \mathcal{M}$ such that $\mu(B) < \infty$ and $\epsilon > 0$ there exists $D \in \mathcal{A}$ such that $\mu(B \triangle D) < \epsilon$.*

Proof. By Corollary 6.38, there exists $C \in \mathcal{A}_\sigma$ such $B \subset C$ and $\mu(C \setminus B) < \epsilon$. Now write $C = \cup_{n=1}^\infty C_n$ with $C_n \in \mathcal{A}$ for each n. By replacing C_n by $\cup_{k=1}^n C_k \in \mathcal{A}$ if necessary, we may assume that $C_n \uparrow C$ as $n \to \infty$. Since $C_n \setminus B \uparrow C \setminus B$ and $B \setminus C_n \downarrow B \setminus C = \emptyset$ as $n \to \infty$ and $\mu(B \setminus C_1) \leq \mu(B) < \infty$, we know that

$$\lim_{n\to\infty} \mu(C_n \setminus B) = \mu(C \setminus B) < \epsilon \text{ and } \lim_{n\to\infty} \mu(B \setminus C_n) = \mu(B \setminus C) = 0$$

Hence for n sufficiently large,

$$\mu(B \triangle C_n) = (\mu(C_n \setminus B) + \mu(B \setminus C_n) < \epsilon.$$

Hence we are done by taking $D = C_n \in \mathcal{A}$ for an n sufficiently large. ■

Remark 6.40. We have to assume that $\mu(B) < \infty$ as the following example shows. Let $X = \mathbb{R}$, $\mathcal{M} = \mathcal{B}$, $\mu = m$, $\mathcal{A}$ be the algebra generated by half open intervals of the form $(a, b]$, and $B = \cup_{n=1}^\infty (2n, 2n+1]$. It is easily checked that for every $D \in \mathcal{A}$, that $m(B \Delta D) = \infty$.

For Exercises 6.1 – 6.3 let $\tau \subseteq \mathcal{P}(X)$ be a topology, $\mathcal{M} = \sigma(\tau)$ and $\mu : \mathcal{M} \to [0, \infty)$ be a finite measure ($\mu(X) < \infty$).

Exercise 6.1. Let

$$\mathcal{F} := \{A \in \mathcal{M} : \mu(A) = \inf\{\mu(V) : A \subseteq V \in \tau\}\}. \tag{6.31}$$

(1) Show $\mathcal{F}$ may be described as the collection of set $A \in \mathcal{M}$ such that for all $\epsilon > 0$ there exists $A \subset V \in \tau$ such that $\mu(V \setminus A) < \epsilon$.
(2) Show $\mathcal{F}$ is a monotone class.

Exercise 6.2. Give an example of a topology τ on $X = \{1, 2\}$ and a measure μ on $\mathcal{M} = \sigma(\tau)$ such that $\mathcal{F}$ defined in Eq. (6.31) is **not** $\mathcal{M}$.

Exercise 6.3. Suppose now $\tau \subseteq \mathcal{P}(X)$ is a topology with the property that to every closed set $C \subset X$, there exists $V_n \in \tau$ such that $V_n \downarrow C$ as $n \to \infty$. Let $\mathcal{A} = \mathcal{A}(\tau)$ be the algebra generated by τ.

(1) With the aid of Exercise 4.1, show that $\mathcal{A} \subset \mathcal{F}$. Therefore by exercise 6.1 and the monotone class theorem, $\mathcal{F} = \mathcal{M}$, i.e.

$$\mu(A) = \inf\{\mu(V) : A \subseteq V \in \tau\}.$$

(**Hint:** Recall the structure of $\mathcal{A}$ from Exercise 4.1.)
(2) Show this result is equivalent to following statement: for every $\epsilon > 0$ and $A \in \mathcal{M}$ there exist a closed set C and an open set V such that $C \subset A \subset V$ and $\mu(V \setminus C) < \epsilon$. (**Hint:** Apply part 1. to both A and A^c.)

Exercise 6.4 (Generalization to the σ – finite case)**.** Let $\tau \subseteq \mathcal{P}(X)$ be a topology with the property that to every closed set $F \subset X$, there exists $V_n \in \tau$ such that $V_n \downarrow F$ as $n \to \infty$. Also let $\mathcal{M} = \sigma(\tau)$ and $\mu : \mathcal{M} \to [0, \infty]$ be a measure which is σ – finite on τ.

(1) Show that for all $\epsilon > 0$ and $A \in \mathcal{M}$ there exists an open set $V \in \tau$ and a closed set F such that $F \subset A \subset V$ and $\mu(V \setminus F) \le \epsilon$.
(2) Let F_σ denote the collection of subsets of X which may be written as a countable union of closed sets. Use item 1. to show for all $B \in \mathcal{M}$, there exists $C \in \tau_\delta$ (τ_δ is customarily written as G_δ) and $A \in F_\sigma$ such that $A \subset B \subset C$ and $\mu(C \setminus A) = 0$.

Exercise 6.5 (Metric Space Examples)**.** Suppose that (X, d) is a metric space and τ_d is the topology of d – open subsets of X. To each set $F \subset X$ and $\epsilon > 0$ let

$$F_\epsilon = \{x \in X : d_F(x) < \epsilon\} = \cup_{x \in F} B_x(\epsilon) \in \tau_d.$$

Show that if F is closed, then $F_\epsilon \downarrow F$ as $\epsilon \downarrow 0$ and in particular $V_n := F_{1/n} \in \tau_d$ are open sets decreasing to F. Therefore the results of Exercises 6.3 and 6.4 apply to measures on metric spaces with the Borel σ – algebra, $\mathcal{B} = \sigma(\tau_d)$.

Corollary 6.41. *Let $\mathcal{B}$ be the Borel σ – algebra on $\mathbb{R}^n$ equipped with the standard topology induced by open balls with respect to the Euclidean distance. Suppose that $\mu : \mathcal{B} \to [0, \infty]$ is a measure such that $\mu(A) < \infty$ whenever A is a bounded set.*

(1) *Then for all $A \in \mathcal{B}$ and $\epsilon > 0$ there exist a closed set F and an open set V such that $F \subset A \subset V$ and $\mu(V \setminus F) < \epsilon$.*
(2) *If $\mu(A) < \infty$, the set F in item 1. may be chosen to be compact.*

(3) *For all* $A \in \mathcal{B}$ *we may compute* $\mu(A)$ *using*

$$\mu(A) = \inf\{\mu(V) : A \subset V \text{ and } V \text{ is open}\} \tag{6.32}$$

$$= \sup\{\mu(K) : K \subset A \text{ and } K \text{ is compact}\}. \tag{6.33}$$

Proof. Item 1. follows from Exercises 6.4 and 6.5. If $\mu(A) < \infty$ and $F \subset A \subset V$ as in item 1. Let

$$K_n := \{x \in F : |x| \leq n\}. \tag{6.34}$$

Then K_n is closed and bounded in $\mathbb{R}^n$ and hence compact and $K_n \uparrow F$ as $n \to \infty$. Since $\mu(A) < \infty$ and $\mu(V \setminus A) < \epsilon$ we know that $\mu(V) < \infty$. Using this fact and the fact that $V \setminus K_n \downarrow V \setminus F$, we conclude that $\mu(V \setminus K_n) \downarrow \mu(V \setminus F) < \epsilon$ as $n \to \infty$. Thus for sufficiently large n we have $K = K_n$ is a compact set such that $K \subset A \subset V$ and $\mu(V \setminus K) < \epsilon$.

Item 1. easily implies that Eq. (6.32) holds and item 2. implies Eq. (6.33) holds when $\mu(A) < \infty$. So we need only check Eq. (6.33) when $\mu(A) = \infty$. By Item 1. there is a closed set $F \subset A$ such that $\mu(A \setminus F) < 1$ and in particular $\mu(F) = \infty$. Letting $K_n \subset F \subset A$ be the compact set as in Eq. (6.34), we have $\mu(K_n) \uparrow \mu(F) = \infty = \mu(A)$ which shows that Eq. (6.33) also holds when $\mu(A) = \infty$. ■

6.6. **Exercises.**

Exercise 6.6. Let $(X_j, \mathcal{M}_j, \mu_j)$ for $j = 1, 2, 3$ be σ – finite measure spaces. Let $F : X_1 \times X_2 \times X_3 \to (X_1 \times X_2) \times X_3$ be defined by

$$F((x_1, x_2), x_3) = (x_1, x_2, x_3).$$

(1) Show F is $((\mathcal{M}_1 \otimes \mathcal{M}_2) \otimes \mathcal{M}_3, \mathcal{M}_1 \otimes \mathcal{M}_2 \otimes \mathcal{M}_3)$ – measurable and F^{-1} is $(\mathcal{M}_1 \otimes \mathcal{M}_2 \otimes \mathcal{M}_3, (\mathcal{M}_1 \otimes \mathcal{M}_2) \otimes \mathcal{M}_3)$ – measurable. That is

$$F : ((X_1 \times X_2) \times X_3, (\mathcal{M}_1 \otimes \mathcal{M}_2) \otimes \mathcal{M}_3) \to (X_1 \times X_2 \times X_3, \mathcal{M}_1 \otimes \mathcal{M}_2 \otimes \mathcal{M}_3)$$

is a "measure theoretic isomorphism."

(2) Let $\lambda := F_* [(\mu_1 \otimes \mu_2) \otimes \mu_3]$, i.e. $\lambda(A) = [(\mu_1 \otimes \mu_2) \otimes \mu_3] (F^{-1}(A))$ for all $A \in \mathcal{M}_1 \otimes \mathcal{M}_2 \otimes \mathcal{M}_3$. Then λ is the unique measure on $\mathcal{M}_1 \otimes \mathcal{M}_2 \otimes \mathcal{M}_3$ such that

$$\lambda(A_1 \times A_2 \times A_3) = \mu_1(A_1)\mu_2(A_2)\mu_3(A_3)$$

for all $A_i \in \mathcal{M}_i$. We will write $\lambda := \mu_1 \otimes \mu_2 \otimes \mu_3$.

(3) Let $f : X_1 \times X_2 \times X_3 \to [0, \infty]$ be a $(\mathcal{M}_1 \otimes \mathcal{M}_2 \otimes \mathcal{M}_3, \mathcal{B}_{\mathbb{R}})$ – measurable function. Verify the identity,

$$\int_{X_1 \times X_2 \times X_3} f d\lambda = \int_{X_3} \int_{X_2} \int_{X_1} f(x_1, x_2, x_3) d\mu_1(x_1) d\mu_2(x_2) d\mu_3(x_3),$$

makes sense and is correct. Also shows that the identity holds for any one of the six possible orderings of the iterated integrals.

Exercise 6.7. Prove the second assertion of Theorem 6.26. That is show m^d is the unique translation invariant measure on $\mathcal{B}_{\mathbb{R}^d}$ such that $m^d((0, 1]^d) = 1$. **Hint:** Look at the proof of Theorem 5.11.

Exercise 6.8. (Part of Folland Problem 2.46 on p. 69.) Let $X = [0,1]$, $\mathcal{M} = \mathcal{B}_{[0,1]}$ be the Borel σ − field on X, m be Lebesgue measure on $[0,1]$ and ν be counting measure, $\nu(A) = \#(A)$. Finally let $D = \{(x,x) \in X^2 : x \in X\}$ be the diagonal in X^2. Show

$$\int_X \int_X 1_D(x,y) d\nu(y) dm(x) \neq \int_X \int_X 1_D(x,y) dm(x) d\nu(y)$$

by explicitly computing both sides of this equation.

Exercise 6.9. Folland Problem 2.48 on p. 69. (Fubini problem.)

Exercise 6.10. Folland Problem 2.50 on p. 69. (Note the $\mathcal{M} \times \mathcal{B}_{\mathbb{R}}$ should be $\mathcal{M} \otimes \mathcal{B}_{\bar{\mathbb{R}}}$ in this problem.)

Exercise 6.11. Folland Problem 2.55 on p. 77. (Explicit integrations.)

Exercise 6.12. Folland Problem 2.56 on p. 77. Let $f \in L^1((0,a), dm)$, $g(x) = \int_x^a \frac{f(t)}{t} dt$ for $x \in (0,a)$, show $g \in L^1((0,a), dm)$ and

$$\int_0^a g(x) dx = \int_0^a f(t) dt.$$

Exercise 6.13. Show $\int_0^\infty \left|\frac{\sin x}{x}\right| dm(x) = \infty$. So $\frac{\sin x}{x} \notin L^1([0,\infty), m)$ and $\int_0^\infty \frac{\sin x}{x} dm(x)$ is not defined as a Lebesgue integral.

Exercise 6.14. Folland Problem 2.57 on p. 77.

Exercise 6.15. Folland Problem 2.58 on p. 77.

Exercise 6.16. Folland Problem 2.60 on p. 77. Properties of Γ − functions.

Exercise 6.17. Folland Problem 2.61 on p. 77. Fractional integration.

Exercise 6.18. Folland Problem 2.62 on p. 80. Rotation invariance of surface measure on S^{n-1}.

Exercise 6.19. Folland Problem 2.64 on p. 80. On the integrability of $|x|^a \left|\log|x|\right|^b$ for x near 0 and x near ∞ in $\mathbb{R}^n$.