CHAPTER 3

Further Construction Methods of Measures

Introduction

In the first section of this chapter we collect some basic results on metric spaces, which every mathematician must know about. Section 3.2 gives a version of the Riesz Representation Theorem, which leads to another and perhaps simpler approach to Lebesgue measure than the Carathéodory Theorem. A reader can skip Section 3.2 without losing the continuity in this paper. The chapter also treats so called product measures and Stieltjes integrals.

3.1. Metric Spaces

The construction of our most important measures requires topological concepts. For our purpose it will be enough to restrict ourselves to so called metric spaces.

A metric d on a set X is a mapping $d : X \times X \to [0, \infty[$ such that

(a) $d(x, y) = 0$ if and only if $x = y$
(b) $d(x, y) = d(y, x)$ (symmetry)
(c) $d(x, y) \leq d(x, z) + d(z, y)$ (triangle inequality).

Here recall, if $A_1, ..., A_n$ are sets,

$$A_1 \times ... \times A_n = \{(x_1, ..., x_n);\ x_i \in A_i \text{ for all } i = 1, ..., n\}$$

A set X equipped with a metric d is called a metric space. Sometimes we write $X = (X, d)$ to emphasize the metric d. If E is a subset of the metric space (X, d), the function $d_{|E \times E}(x, y) = d(x, y)$, if $x, y \in E$, is a metric on E. Thus $(E, d_{|E \times E})$ is a metric space.

The function $\varphi(t) = \min(1, t)$, $t \geq 0$, satisfies the inequality

$$\varphi(s + t) \leq \varphi(s) + \varphi(t).$$

Therefore, if d is a metric on X, $\min(1, d)$ is a metric on X. The metric $\min(1, d)$ is a bounded metric.

The set $\mathbf{R}$ equipped with the metric $d_1(x, y) = \mid x - y \mid$ is a metric space. More generally, $\mathbf{R}^n$ equipped with the metric

$$d_n(x, y) = d_n((x_1, ..., x_n), (y_1, ..., y_n)) = \max_{1 \leq k \leq n} \mid x_k - y_k \mid$$

is a metric space. If not otherwise stated, it will always be assumed that $\mathbf{R}^n$ is equipped with this metric.

Let $C\,[0, T]$ denote the vector space of all real-valued continuous functions on the interval $[0, T]$, where $T > 0$. Then

$$d_\infty(x, y) = \max_{0 \leq t \leq T} \mid x(t) - y(t) \mid$$

is a metric on $C\,[0, T]$.

If (X_k, e_k), $k = 1, ..., n$, are metric spaces,

$$d(x, y) = \max_{1 \leq k \leq n} e_k(x_k, y_k),\ \ x = (x_1, ..., x_n)\ , y = (y_1, ..., y_n)$$

is a metric on $X_1 \times ... \times X_n$. The metric d is called the product metric on $X_1 \times ... \times X_n$.

If $X = (X, d)$ is a metric space and $x \in X$ and $r > 0$, the open ball with centre at x and radius r is the set $B(x, r) = \{y \in X; d(y, x) < r\}$. If $E \subseteq X$ and E is contained in an appropriate open ball in X it is said to be bounded. The diameter of E is, by definition,

$$\text{diam } E = \sup_{x, y \in E} d(x, y)$$

and it follows that E is bounded if and only if diam $E < \infty$. A subset of X which is a union of open balls in X is called open. In particular, an open ball is an open set. The empty set is open since the union of an empty family of sets is empty. An arbitrary union of open sets is open. The class of all open subsets of X is called the topology of X. The metrics d and $\min(1,d)$ determine the same topology. A subset E of X is said to be closed if its complement E^c relative to X is open. An intersection of closed subsets of X is closed. If $E \subseteq X$, E° denotes the largest open set contained in E and E^- (or $\bar{E}$) the smallest closed set containing E. E° is the interior of E and E^- its closure. The σ-algebra generated by the open sets in X is called the Borel σ-algebra in X and is denoted by $\mathcal{B}(X)$. A positive measure on $\mathcal{B}(X)$ is called a positive Borel measure.

A sequence $(x_n)_{n=1}^\infty$ in X converges to $x \in X$ if

$$\lim_{n\to\infty} d(x_n, x) = 0.$$

If, in addition, the sequence $(x_n)_{n=1}^\infty$ converges to $y \in X$, the inequalities

$$0 \leq d(x,y) \leq d(x_n,x) + d(x_n,y)$$

imply that $y = x$ and the limit point x is unique.

If $E \subseteq X$ and $x \in X$, the following properties are equivalent:

(i) $x \in E^-$.
(ii) $B(x,r) \cap E \neq \phi$, all $r > 0$.
(iii) There is a sequence $(x_n)_{n=1}^\infty$ in E which converges to x.

If $B(x,r) \cap E = \phi$, then $B(x,r)^c$ is a closed set containing E but not x. Thus $x \notin E^-$. This proves that (i)$\Rightarrow$(ii). Conversely, if $x \notin E^-$, since $\bar{E}^c$ is open there exists an open ball $B(y,s)$ such that $x \in B(y,s) \subseteq \bar{E}^c \subseteq E^c$. Now choose $r = s - d(x,y) > 0$ so that $B(x,r) \subseteq B(y,s)$. Then $B(x,r) \cap E = \phi$. This proves (ii)$\Rightarrow$(i).

If (ii) holds choose for each $n \in \mathbf{N}_+$ a point $x_n \in E$ with $d(x_n,x) < \frac{1}{n}$ and (iii) follows. If there exists an $r > 0$ such that $B(x,r) \cap E = \phi$, then (iii) cannot hold. Thus (iii)$\Rightarrow$(ii).

If $E \subseteq X$, the set $E^- \setminus E^\circ$ is called the boundary of E and is denoted by ∂E.

A set $A \subseteq X$ is said to be dense in X if $A^- = X$. The metric space X is called separable if there is an at most denumerable dense subset of X. For example, $\mathbf{Q}^n$ is a dense subset of $\mathbf{R}^n$. The space $\mathbf{R}^n$ is separable.

Theorem 3.1.1. $\mathcal{B}(\mathbf{R}^n) = \mathcal{R}_n$.

PROOF. The σ-algebra $\mathcal{R}_n$ is generated by the open n-cells in $\mathbf{R}^n$ and an open n-cell is an open subset of $\mathbf{R}^n$. Hence $\mathcal{R}_n \subseteq \mathcal{B}(\mathbf{R}^n)$. Let U be an open subset in $\mathbf{R}^n$ and note that an open ball in $\mathbf{R}^n = (\mathbf{R}^n, d_n)$ is an open n-cell. If $x \in U$ there exist an $a \in \mathbf{Q}^n \cap U$ and a rational number $r > 0$ such that $x \in B(a, r) \subseteq U$. Thus U is an at most denumerable union of open n-cells and it follows that $U \in \mathcal{R}_n$. Thus $\mathcal{B}(\mathbf{R}^n) \subseteq \mathcal{R}_n$ and the theorem is proved.

Let $X = (X, d)$ and $Y = (Y, e)$ be two metric spaces. A mapping $f : X \to Y$ (or $f : (X, d) \to (Y, e)$ to emphasize the underlying metrics) is said to be continuous at the point $a \in X$ if for every $\varepsilon > 0$ there exists a $\delta > 0$ such that

$$x \in B(a, \delta) \Rightarrow f(x) \in B(f(a), \varepsilon).$$

Equivalently this means that for any sequence $(x_n)_{n=1}^{\infty}$ in X which converges to a in X, the sequence $(f(x_n))_{n=1}^{\infty}$ converges to $f(a)$ in Y. If f is continuous at each point of X, the mapping f is called continuous. Stated otherwise this means that

$$f^{-1}(V) \text{ is open if } V \text{ is open}$$

or

$$f^{-1}(F) \text{ is closed if } F \text{ is closed.}$$

The mapping f is said to be Borel measurable if

$$f^{-1}(B) \in \mathcal{B}(X) \text{ if } B \in \mathcal{B}(Y)$$

or, what amounts to the same thing,

$$f^{-1}(V) \in \mathcal{B}(X) \text{ if } V \text{ is open.}$$

A Borel measurable function is sometimes called a Borel function. A continuous function is a Borel function.

Example 3.1.1. Let $f : (\mathbf{R}, d_1) \to (\mathbf{R}, d_1)$ be a continuous strictly increasing function and set $\rho(x, y) = \mid f(x) - f(y) \mid$, $x, y \in \mathbf{R}$. Then ρ is a metric on $\mathbf{R}$.

Define $j(x) = x$, $x \in \mathbf{R}$. The mapping $j : (\mathbf{R},d_1) \rightarrow (\mathbf{R},\rho)$ is continuous. We claim that the map $j : (\mathbf{R},\rho) \rightarrow (\mathbf{R},d_1)$ is continuous. To see this, let $a \in \mathbf{R}$ and suppose the sequence $(x_n)_{n=1}^{\infty}$ converges to a in the metric space $(\mathbf{R},\rho)$, that is $\mid f(x_n) - f(a) \mid \rightarrow 0$ as $n \rightarrow \infty$. Let $\varepsilon > 0$. Then

$$f(x_n) - f(a) \geq f(a+\varepsilon) - f(a) > 0 \text{ if } x_n \geq a + \varepsilon$$

and

$$f(a) - f(x_n) \geq f(a) - f(a-\varepsilon) > 0 \text{ if } x_n \leq a - \varepsilon.$$

Thus $x_n \in \,]a-\varepsilon, a+\varepsilon[$ if n is sufficiently large. This proves that he map $j : (\mathbf{R},\rho) \rightarrow (\mathbf{R},d_1)$ is continuous.

The metrics d_1 and ρ determine the same topology and Borel subsets of $\mathbf{R}$.

A mapping $f : (X,d) \rightarrow (Y,e)$ is said to be uniformly continuous if for each $\varepsilon > 0$ there exists a $\delta > 0$ such that $e(f(x),f(y)) < \varepsilon$ as soon as $d(x,y) < \delta$.

If $x \in X$ and E, $F \subseteq X$, let

$$d(x,E) = \inf_{u \in E} d(x,u)$$

be the distance from x to E and let

$$d(E,F) = \inf_{u \in E, v \in F} d(u,v)$$

be the distance between E and F. Note that $d(x,E) = 0$ if and only if $x \in \bar{E}$.

If $x, y \in X$ and $u \in E$,

$$d(x,u) \leq d(x,y) + d(y,u)$$

and, hence

$$d(x,E) \leq d(x,y) + d(y,u)$$

and

$$d(x,E) \leq d(x,y) + d(y,E).$$

Next suppose $E \neq \phi$. Then by interchanging the roles of x and y, we get

$$\mid d(x,E) - d(y,E) \mid \leq d(x,y)$$

and conclude that the distance function $d(x, E)$, $x \in X$, is continuous. In fact, it is uniformly continuous. If $x \in X$ and $r > 0$, the so called closed ball $\bar{B}(x, r) = \{y \in X;\ d(y, x) \leq r\}$ is a closed set since the map $y \to d(y, x)$, $y \in X$, is continuous.

If $F \subseteq X$ is closed and $\varepsilon > 0$, the continuous function

$$\Pi^X_{F,\varepsilon} = \max(0, 1 - \frac{1}{\varepsilon} d(\cdot, F))$$

fulfils $0 \leq \Pi^X_{F,\varepsilon} \leq 1$ and $\Pi^X_{F,\varepsilon} = 1$ on F. Furthermore, $\Pi^X_{F,\varepsilon}(a) > 0$ if and only if $a \in F_\varepsilon =_{def} \{x \in X;\ d(x, F) < \varepsilon\}$. Thus

$$\chi_F \leq \Pi^X_{F,\varepsilon} \leq \chi_{F_\varepsilon}.$$

Let $X = (X, d)$ be a metric space. A sequence $(x_n)_{n=1}^\infty$ in X is called a Cauchy sequence if to each $\varepsilon > 0$ there exists a positive integer p such that $d(x_n, x_m) < \varepsilon$ for all $n, m \geq p$. If a Cauchy sequence $(x_n)_{n=1}^\infty$ contains a convergent subsequence $(x_{n_k})_{k=1}^\infty$ it must be convergent. To prove this claim, suppose the subsequence $(x_{n_k})_{k=1}^\infty$ converges to a point $x \in X$. Then

$$d(x_m, x) \leq d(x_m, x_{n_k}) + d(x_{n_k}, x)$$

can be made arbitrarily small for all sufficiently large m by choosing k sufficiently large. Thus $(x_n)_{n=1}^\infty$ converges to x.

A subset E of X is said to be complete if every Cauchy sequence in E converges to a point in E. If $E \subseteq X$ is closed and X is complete it is clear that E is complete. Conversely, if X is a metric space and a subset E of X is complete, then E is closed.

It is important to know that $\mathbf{R}$ is complete equipped with its standard metric. To see this let $(x_n)_{n=1}^\infty$ be a Cauchy sequence. There exists a positive integer such that $\mid x_n - x_m \mid < 1$ if $n, m \geq p$. Therefore

$$\mid x_n \mid \leq \mid x_n - x_p \mid + \mid x_p \mid \leq 1 + \mid x_p \mid$$

for all $n \geq p$. We have proved that the sequence $(x_n)_{n=1}^\infty$ is bounded (the reader can check that every Cauchy sequence in a metric space has this property). Now define

$$a = \sup \{x \in \mathbf{R};\ \text{there are only finitely many } n \text{ with } x_n \leq x\}.$$

The definition implies that there exists a subsequence $(x_{n_k})_{k=1}^\infty$, which converges to a (since for any $r > 0$, $x_n \in B(a, r)$ for infinitely many n). The original sequence is therefore convergent and we conclude that $\mathbf{R}$ is complete (equipped with its standard metric d_1). It is simple to prove that the product of n complete spaces is complete and we conclude that $\mathbf{R}^n$ is complete.

Let $E \subseteq X$. A family $(V_i)_{i\in I}$ of subsets of X is said to be a cover of E if $\cup_{i\in I} V_i \supseteq E$ and E is said to be covered by the $V_i's$. The cover $(V_i)_{i\in I}$ is said to be an open cover if each member V_i is open. The set E is said to be totally bounded if, for every $\varepsilon > 0$, E can be covered by finitely many open balls of radius ε. A subset of a totally bounded set is totally bounded.

The following definition is especially important.

Definition 3.1.1. A subset E of a metric space X is said to be compact if to every open cover $(V_i)_{i\in I}$ of E, there is a finite subcover of E, which means there is a finite subset J of I such that $(V_i)_{i\in J}$ is a cover of E.

If K is closed, $K \subseteq E$, and E is compact, then K is compact. To see this, let $(V_i)_{i\in I}$ be an open cover of K. This cover, augmented by the set $X \setminus K$ is an open cover of E and has a finite subcover since E is compact. Noting that $K \cap (X \setminus K) = \phi$, the assertion follows.

Theorem 3.1.2. *The following conditions are equivalent:*

(a) *E is complete and totally bounded.*

(b) *Every sequence in E contains a subsequence which converges to a point of E.*

(c) *E is compact.*

PROOF. (a)$\Rightarrow$(b). Suppose $(x_n)_{n=1}^{\infty}$ is a sequence in E. The set E can be covered by finitely many open balls of radius 2^{-1} and at least one of them must contain x_n for infinitely many $n \in \mathbf{N}_+$. Suppose $x_n \in B(a_1, 2^{-1})$ if $n \in N_1 \subseteq N_0 =_{def} \mathbf{N}_+$, where N_1 is infinite. Next $E \cap B(a_1, 2^{-1})$ can be covered by finitely many balls of radius 2^{-2} and at least one of them must contain x_n for infinitely many $n \in N_1$. Suppose $x_n \in B(a_2, 2^{-1})$ if $n \in N_2$, where $N_2 \subseteq N_1$ is infinite. By induction, we get open balls $B(a_j, 2^{-j})$ and infinite sets $N_j \subseteq N_{j-1}$ such that $x_n \in B(a_j, 2^{-j})$ for all $n \in N_j$ and $j \geq 1$.

Let $n_1 < n_2 < ...$, where $n_k \in N_k$, $k = 1, 2, ...$. The sequence $(x_{n_k})_{k=1}^{\infty}$ is a Cauchy sequence, and since E is complete it converges to a point of E .

(b)⇒(a). If E is not complete there is a Cauchy sequence in E with no limit in E. Therefore no subsequence can converge in E, which contradicts (b). On the other hand if E is not totally bounded, there is an $\varepsilon > 0$ such that E cannot be covered by finitely many balls of radius ε. Let $x_1 \in E$ be arbitrary. Having chosen $x_1, ..., x_{n-1}$, pick $x_n \in E\setminus \cup_{i=1}^{n-1} B(x_i, \varepsilon)$, and so on. The sequence $(x_n)_{n=1}^{\infty}$ cannot contain any convergent subsequence as $d(x_n, x_m) \geq \varepsilon$ if $n \neq m$, which contradicts (b).

{(a) and (b)} ⇒(c). Let $(V_i)_{i\in I}$ be an open cover of E. Since E is totally bounded it is enough to show that there is an $\varepsilon > 0$ such that any open ball of radius ε which intersects E is contained in some V_i. Suppose on the contrary that for every $n \in \mathbf{N}_+$ there is an open ball B_n of radius $\leq 2^{-n}$ which intersects E and is contained in no V_i. Choose $x_n \in B_n \cap E$ and assume without loss of generality that $(x_n)_{n=1}^{\infty}$ converges to some point x in E by eventually going to a subsequence. Suppose $x \in V_{i_0}$ and choose $r > 0$ such that $B(x, r) \subseteq V_{i_0}$. But then $B_n \subseteq B(x, r) \subseteq V_{i_0}$ for large n, which contradicts the assumption on B_n.

(c)⇒(b). If $(x_n)_{n=1}^{\infty}$ is a sequence in E with no convergent subsequence in E, then for every $x \in E$ there is an open ball $B(x, r_x)$ which contains x_n for only finitely many n. Then $(B(x, r_x))_{x\in E}$ is an open cover of E without a finite subcover.

Corollary 3.1.1. *A subset of* $\mathbf{R}^n$ *is compact if and only if it is closed and bounded.*

PROOF. Suppose K is compact. If $x_n \in K$ and $x_n \notin B(0, n)$ for every $n \in \mathbf{N}_+$, the sequence $(x_n)_{n=1}^{\infty}$ cannot contain a convergent subsequence. Thus K is bounded. Since K is complete it is closed.

Conversely, suppose K is closed and bounded. Since $\mathbf{R}^n$ is complete and K is closed, K is complete. We next prove that a bounded set is totally bounded. It is enough to prove that any n-cell in $\mathbf{R}^n$ is a union of finitely many n-cells $I_1 \times ... \times I_n$ where each interval $I_1, ..., I_n$ has a prescribed positive length. This is clear and the theorem is proved.

Corollary 3.1.2. *Suppose $f : X \to \mathbf{R}$ is continuous and X compact.*
(a) *There exists an $a \in X$ such that $\max_X f = f(a)$ and a $b \in X$ such that $\min_X f = f(b)$.*
(b) *The function f is uniformly continuous.*

PROOF. (a) For each $a \in X$, let $V_a = \{x \in X : f(x) < 1 + f(a)\}$. The open cover $(V_a)_{a\in K}$ of X has a finite subcover and it follows that f is bounded. Let $(x_n)_{n=1}^\infty$ be a sequence in X such that $f(x_n) \to \sup_K f$ as $n \to \infty$. Since X is compact there is a subsequence $(x_{n_k})_{k=1}^\infty$ which converges to a point $a \in X$. Thus, by the continuity of f, $f(x_{n_k}) \to f(a)$ as $k \to \infty$.

The existence of a minimum is proved in a similar way.

(b) If f is not uniformly continuous there exist $\varepsilon > 0$ and sequences $(x_n)_{n=1}^\infty$ and $(y_n)_{n=1}^\infty$ such that $\mid f(x_n) - f(y_n) \mid \geq \varepsilon$ and $\mid x_n - y_n \mid < 2^{-n}$ for every $n \geq 1$. Since X is compact there exists a subsequence $(x_{n_k})_{k=1}^\infty$ of $(x_n)_{n=1}^\infty$ which converges to a point $a \in X$. Clearly the sequence $(y_{n_k})_{k=1}^\infty$ converges to a and therefore

$$\mid f(x_{n_k}) - f(y_{n_k}) \mid \leq \mid f(x_{n_k}) - f(a) \mid + \mid f(a) - f(y_{n_k}) \mid \to 0$$

as $k \to \infty$ since f is continuous. But $\mid f(x_{n_k}) - f(y_{n_k}) \mid \geq \varepsilon$ and we have got a contradiction. The corollary is proved.

Example 3.1.2. Suppose $X =]0,1]$ and define $\rho_1(x,y) = d_1(x,y)$ and $\rho_2(x,y) = \mid \frac{1}{x} - \frac{1}{y} \mid$, $x, y \in X$. As in Example 3.1.1 we conclude that the metrics ρ_1 and ρ_2 determine the same topology of subsets of X. The space (X, ρ_1)

totally bounded but not complete. However, the space (X,ρ_2) is not totally bounded but it is complete. To see this, let $(x_n)_{n=1}^{\infty}$be a Cauchy sequence in (X,ρ_2). As a Cauchy sequence it must be bounded and therefore there exists an $\varepsilon \in\ ]0,1]$ such that $x_n \in [\varepsilon,1]$ for all n. But then, by Corollary 3.1.1, $(x_n)_{n=1}^{\infty}$ contains a convergent subsequence in (X,ρ_1) and, accordingly from this, the same property holds in (X,ρ_2). The space (X,ρ_2) is not compact, since (X,ρ_1) is not compact, and we conclude from Theorem 3.1.2 that the space (X,ρ_2) cannot be totally bounded.

Example 3.1.3. Set $\hat{\mathbf{R}}$=$\mathbf{R}\cup\{-\infty,\infty\}$ and

$$\hat{d}(x,y) =\mid \arctan x - \arctan y \mid$$

if $x,y\in\hat{\mathbf{R}}$. Here

$$\arctan\infty = \frac{\pi}{2} \text{ and } \arctan-\infty = -\frac{\pi}{2}.$$

Example 3.1.1 shows that the standard metric d_1 and the metric $\hat{d}_{|\mathbf{R}\times\mathbf{R}}$ determine the same topology.

We next prove that the metric space $\hat{\mathbf{R}}$ is compact. To this end, consider a sequence $(x_n)_{n=1}^{\infty}$ in $\hat{\mathbf{R}}$. If there exists a real number M such that $\mid x_n \mid\leq M$ for infinitely many n, the sequence $(x_n)_{n=1}^{\infty}$ contains a convergent subsequence since the interval $[-M,M]$ is compact. In the opposite case, for each positive real number M, either $x_n\geq M$ for infinitely many n or $x_n\leq -M$ for infinitely many n. Suppose $x_n\geq M$ for infinitely many n for every $M\in\mathbf{N}_+$. Then $\hat{d}(x_{n_k},\infty) =\mid \arctan x_{n_k} - \frac{\pi}{2} \mid\to 0$ as $k\to\infty$ for an appropriate subsequence $(x_{n_k})_{k=1}^{\infty}$.

The space $\hat{\mathbf{R}}$= $(\hat{\mathbf{R}},\hat{d})$ is called a two-point compactification of $\mathbf{R}$.

It is an immediate consequence of Theorem 3.1.2 that the product of finitely many compact metric spaces is compact. Thus $\hat{\mathbf{R}}^n$ equipped with the product metric is compact.

We will finish this section with several useful approximation theorems.

Theorem 3.1.3. *Suppose X is a metric space and μ positive Borel measure in X. Moreover, suppose there is a sequence $(U_n)_{n=1}^{\infty}$of open subsets of X such that*

$$X=\cup_{n=1}^{\infty}U_n$$

and

$$\mu(U_n) < \infty, \text{ all } n \in \mathbf{N}_+.$$

Then for each $A \in \mathcal{B}(X)$ *and* $\varepsilon > 0$, *there are a closed set* $F \subseteq A$ *and an open set* $V \supseteq A$ *such that*

$$\mu(V \setminus F) < \varepsilon.$$

In particular, for every $A \in \mathcal{B}(X)$,

$$\mu(A) = \inf_{\substack{V \supseteq A \\ V \text{ open}}} \mu(V)$$

and

$$\mu(A) = \sup_{\substack{F \subseteq A \\ F \text{ closed}}} \mu(F)$$

If $X = \mathbf{R}$ and $\mu(A) = \Sigma_{n=1}^{\infty} \delta_{\frac{1}{n}}(A)$, $A \in \mathcal{R}$, then $\mu(\{0\}) = 0$ and $\mu(V) = \infty$ for every open set containing $\{0\}$. The hypothesis that the sets U_n, $n \in \mathbf{N}_+$, are open (and not merely Borel sets) is very important in Theorem 3.1.3.

PROOF. First suppose that μ is a finite positive measure.

Let $\mathcal{A}$ be the class of all Borel sets A in X such that for every $\varepsilon > 0$ there exist a closed $F \subseteq A$ and an open $V \supseteq A$ such that $\mu(V \setminus F) < \varepsilon$. If F is a closed subset of X and $V_n = \left\{x; d(x, F) < \frac{1}{n}\right\}$, then V_n is open and, by Theorem 1.1.2 (f), $\mu(V_n) \downarrow v(F)$ as $n \to \infty$. Thus $F \in \mathcal{A}$ and we conclude that $\mathcal{A}$ contains all closed subsets of X.

Now suppose $A \in \mathcal{A}$. We will prove that $A^c \in \mathcal{A}$. To this end, we choose $\varepsilon > 0$ and a closed set $F \subseteq A$ and an open set $V \supseteq A$ such that $\mu(V \setminus F) < \varepsilon$. Then $V^c \subseteq A^c \subseteq F^c$ and, moreover, $\mu(F^c \setminus V^c) < \varepsilon$ since

$$V \setminus F = F^c \setminus V^c.$$

If we note that V^c is closed and F^c open it follows that $A^c \in \mathcal{A}$.

Next let $(A_i)_{i=1}^{\infty}$ be a denumerable collection of members of $\mathcal{A}$. Choose $\varepsilon > 0$. By definition, for each $i \in \mathbf{N}_+$ there exist a closed $F_i \subseteq A_i$ and an open $V_i \supseteq A_i$ such that $\mu(V_i \setminus F_i) < 2^{-i}\varepsilon$. Set

$$V = \cup_{i=1}^{\infty} V_i.$$

Then

$$\mu(V \setminus (\cup_{i=1}^{\infty} F_i)) \leq \mu(\cup_{i=1}^{\infty}(V_i \setminus F_i))$$
$$\leq \Sigma_{i=1}^{\infty} \mu(V_i \setminus F_i) < \varepsilon.$$

But

$$V \setminus (\cup_{i=1}^{\infty} F_i) = \cap_{n=1}^{\infty} \{V \setminus (\cup_{i=1}^{n} F_i)\}$$

and since μ is a finite positive measure

$$\mu(V \setminus (\cup_{i=1}^{\infty} F_i)) = \lim_{n\to\infty} \mu(V \setminus (\cup_{i=1}^{n} F_i)).$$

Accordingly, from these equations

$$\mu(V \setminus (\cup_{i=1}^{n} F_i)) < \varepsilon$$

if n is large enough. Since a union of open sets is open and a finite union of closed sets is closed, we conclude that $\cup_{i=1}^{\infty} A_i \in \mathcal{A}$. This proves that $\mathcal{A}$ is a σ-algebra. Since $\mathcal{A}$ contains each closed subset of X, $\mathcal{A} = \mathcal{B}(X)$.

We now prove the general case. Suppose $A \in \mathcal{B}(X)$. Since μ^{U_n} is a finite positive measure the previous theorem gives us an open set $V_n \supseteq A \cap U_n$ such that $\mu^{U_n}(V_n \setminus (A \cap U_n)) < \varepsilon 2^{-n}$. By eventually replacing V_n by $V_n \cap U_n$ we can assume that $V_n \subseteq U_n$. But then $\mu(V_n \setminus (A \cap U_n)) = \mu^{U_n}(V_n \setminus (A \cap U_n)) < \varepsilon 2^{-n}$.

Set $V = \cup_{n=1}^{\infty} V_n$ and note that V is open. Moreover,

$$V \setminus A \subseteq \cup_{n=1}^{\infty}(V_n \setminus (A \cap U_n))$$

and we get

$$\mu(V \setminus A) \leq \Sigma_{n=1}^{\infty} \mu(V_n \setminus (A \cap U_n)) < \varepsilon.$$

By applying the result already proved to the complement A^c we conclude there exists an open set $W \supseteq A^c$ such that

$$\mu(A \setminus W^c) = \mu(W \setminus A^c) < \varepsilon.$$

Thus if $F =_{def} W^c$ it follows that $F \subseteq A \subseteq V$ and $\mu(V \setminus F) < 2\varepsilon$. The theorem is proved.

If X is a metric space $C(X)$ denotes the vector space of all real-valued continuous functions $f : X \to \mathbf{R}$. If $f \in C(X)$, the closure of the set of all x where $f(x) \neq 0$ is called the support of f and is denoted by supp$f$. The vector space of all all real-valued continuous functions $f : X \to \mathbf{R}$ with compact support is denoted by $C_c(X)$.

Corollary 3.1.3. *Suppose μ and ν are positive Borel measures in $\mathbf{R}^n$ such that*

$$\mu(K) < \infty \text{ and } \nu(K) < \infty$$

for every compact subset K of $\mathbf{R}^n$. If

$$\int_{\mathbf{R}^n} f(x)d\mu(x) = \int_{\mathbf{R}^n} f(x)d\nu(x), \text{ all } f \in C_c(\mathbf{R}^n)$$

then $\mu = \nu$.

PROOF. Let F be closed. Clearly $\mu(B(0,i)) < \infty$ and $\nu(B(0,i)) < \infty$ for every positive integer i. Hence, by Theorem 3.1.3 it is enough to show that $\mu(F) = \nu(F)$. Now fix a positive integer i and set $K = \bar{B}(0,i) \cap F$. It is enough to show that $\mu(K) = \nu(K)$. But

$$\int_{\mathbf{R}^n} \Pi^{\mathbf{R}^n}_{K,2^{-j}}(x)d\mu(x) = \int_{\mathbf{R}^n} \Pi^{\mathbf{R}^n}_{K,2^{-j}}(x)d\nu(x)$$

for each positive integer j and letting $j \to \infty$ we are done.

A metric space X is called a standard space if it is separable and complete. Standard spaces have a series of very nice properties related to measure theory; an example is furnished by the following

Theorem 3.1.4. (Ulam's Theorem) *Let X be a standard space and suppose μ is a finite positive Borel measure on X. Then to each $A \in \mathcal{B}(X)$ and $\varepsilon > 0$ there exist a compact $K \subseteq A$ and an open $V \supseteq A$ such that $\mu(V \setminus K) < \varepsilon$.*

PROOF. Let $\varepsilon > 0$. We first prove that there is a compact subset K of X such that $\mu(K) > \mu(X) - \varepsilon$. To this end, let A be a dense denumerable subset of X and let $(a_i)_{i=1}^{\infty}$ be an enumeration of A. Now for each positive integer j, $\cup_{i=1}^{\infty} B(a_i, 2^{-j}\varepsilon) = X$, and therefore there is a positive integer n_j such that

$$\mu(\cup_{i=1}^{n_j} B(a_i, 2^{-j}\varepsilon)) > \mu(X) - 2^{-j}\varepsilon.$$

Set

$$F_j = \cup_{i=1}^{n_j} \bar{B}(a_i, 2^{-j}\varepsilon)$$

and

$$L = \cap_{j=1}^{\infty} F_j.$$

The set L is totally bounded. Since X is complete and L closed, L is complete. Therefore, the set L is compact and, moreover

$$\mu(K) = \mu(X) - \mu(L^c) = \mu(X) - \mu(\cup_{j=1}^{\infty} F_j^c)$$

$$\geq \mu(X) - \Sigma_{j=1}^{\infty} \mu(F_j^c) = \mu(X) - \Sigma_{j=1}^{\infty} (\mu(X) - \mu(F_j))$$

$$\geq \mu(X) - \Sigma_{j=1}^{\infty} 2^{-j}\varepsilon = \mu(X) - \varepsilon.$$

Depending on Theorem 3.1.3 to each $A \in \mathcal{B}(X)$ there exists a closed $F \subseteq A$ and an open $V \supseteq A$ such that $\mu(V \setminus F) < \varepsilon$. But

$$V \setminus (F \cap L) = (V \setminus F) \cup (F \setminus L)$$

and we get

$$\mu(V \setminus (F \cap L)) \leq \mu(V \setminus F) + \mu(X \setminus K) < 2\varepsilon.$$

Since the set $F \cap L$ is compact Theorem 3.1.4 is proved.

Two Borel sets in $\mathbf{R}^n$ are said to be almost disjoint if their intersection has volume measure zero.

Theorem 3.1.5. *Every open set U in $\mathbf{R}^n$ is the union of an at most denumerable collection of mutually almost disjoint cubes.*

Before the proof observe that a cube in $\mathbf{R}^n$ is the same as a closed ball in $\mathbf{R}^n$ equipped with the metric d_n.

PROOF. For each, $k \in \mathbf{N}_+$, let $\mathcal{Q}_k$be the class of all cubes of side length 2^{-k} whose vertices have coordinates of the form $i2^{-k}$, $i \in \mathbf{Z}$. Let F_1 be the union of those cubes in $\mathcal{Q}_1$ which are contained in U. Inductively, for $k \geq 1$, let F_k be the union of those cubes in $\mathcal{Q}_k$ which are contained in U and whose interiors are disjoint from $\cup_{j=1}^{k-1} F_j$. Since $d(x, \mathbf{R}^n \setminus U) > 0$ for every $x \in U$ it follows that $U = \cup_{j=1}^{\infty} F_j$.

Exercises

1. Suppose $f : (X, \mathcal{M}) \to (\mathbf{R}^d, \mathcal{R}_d)$ and $g : (X, \mathcal{M}) \to (\mathbf{R}^n, \mathcal{R}_n)$ are measurable. Set $h(x) = (f(x), g(x)) \in \mathbf{R}^{d+n}$ if $x \in X$. Prove that $h : (X, \mathcal{M}) \to (\mathbf{R}^{d+n}, \mathcal{R}_{d+n})$ is measurable.

2. Suppose $f : (X, \mathcal{M}) \to (\mathbf{R}, \mathcal{R})$ and $g : (X, \mathcal{M}) \to (\mathbf{R}, \mathcal{R})$ are measurable. Prove that fg is $(\mathcal{M}, \mathcal{R})$-measurable.

3. The function $f : \mathbf{R} \to \mathbf{R}$ is a Borel function. Set $g(x, y) = f(x)$, $(x, y) \in \mathbf{R}^2$. Prove that $g : \mathbf{R}^2 \to \mathbf{R}$ is a Borel function.

4. Suppose $f : [0, 1] \to \mathbf{R}$ is a continuous function and $g : [0, 1] \to [0, 1]$ a Borel function. Compute the limit

$$\lim_{n \to \infty} \int_0^1 f(g(x)^n) dx.$$

5. Suppose X and Y are metric spaces and $f : X \to Y$ a continuous mapping. Show that $f(E)$ is compact if E is a compact subset of X.

6. Suppose X and Y are metric spaces and $f : X \to Y$ a continuous bijection. Show that the inverse mapping f^{-1} is continuous if X is compact.

7. Construct an open bounded subset V of $\mathbf{R}$ such that $m(\partial V) > 0$.

8. The function $f : [0,1] \to \mathbf{R}$ has a continuous derivative. Prove that the set $f(K) \in \mathcal{Z}_m$ if $K = (f')^{-1}(\{0\})$.

9. Let P denote the class of all Borel probability measures on $[0,1]$ and L the class of all functions $f : [0,1] \to [-1,1]$ such that

$$\mid f(x) - f(y) \mid \leq \mid x - y \mid, \ x, y \in [0,1] \, .$$

For any $\mu, \nu \in P$, define

$$\rho(\mu, \nu) = \sup_{f \in L} \mid \int_{[0,1]} f d\mu - \int_{[0,1]} f d\nu \mid \, .$$

(a) Show that (P, ρ) is a metric space. (b) Compute $\rho(\mu, \nu)$ if μ is linear measure on $[0,1]$ and $\nu = \frac{1}{n}\Sigma_{k=0}^{n-1}\delta_{\frac{k}{n}}$, where $n \in \mathbf{N}_+$ (linear measure on $[0,1]$ is Lebesgue measure on $[0,1]$ restricted to the Borel sets in $[0,1]$).

10. Suppose μ is a finite positive Borel measure on $\mathbf{R}^n$. (a) Let $(V_i)_{i \in I}$ be a family of open subsets of $\mathbf{R}^n$ and $V = \cup_{i \in I} V_i$. Prove that

$$\mu(V) = \sup_{\substack{i_1,...,i_k \in I \\ k \in \mathbf{N}_+}} \mu(V_{i_1} \cup ... \cup V_{i_k}).$$

(b) Let $(F_i)_{i \in I}$ be a family of closed subsets of $\mathbf{R}^n$ and $F = \cap_{i \in I} F_i$. Prove that

$$\mu(F) = \inf_{\substack{i_1,...,i_k \in I \\ k \in \mathbf{N}_+}} \mu(F_{i_1} \cap ... \cap F_{i_k}).$$

$\downarrow\downarrow\downarrow$

3.2. Linear Functionals and Measures

Let X be a metric space. A mapping $T : C_c(X) \to \mathbf{R}$ is said to be a linear functional on $C_c(X)$ if

$$T(f+g) = Tf + Tg, \text{ all } f, g \in C_c(X)$$

and

$$T(\alpha f) = \alpha T f, \text{ all } \alpha \in \mathbf{R}, \ f \in C_c(X).$$

If in addition $Tf \geq 0$ for all $f \geq 0$, T is called a positive linear functional on $C_c(X)$. In this case $Tf \leq Tg$ if $f \leq g$ since $g - f \geq 0$ and $Tg - Tf = T(g-f) \geq 0$. Note that $C_c(X) = C(X)$ if X is compact.

The main result in this section is the following

Theorem 3.2.1. (The Riesz Representation Theorem) *Suppose X is a compact metric space and let T be a positive linear functional on $C(X)$. Then there exists a unique finite positive Borel measure μ in X with the following properties:*

(a)

$$Tf = \int_X f d\mu, \ f \in C(X).$$

(b) *For every $E \in \mathcal{B}(X)$*

$$\mu(E) = \sup_{\substack{K \subseteq E \\ K \text{ compact}}} \mu(K).$$

(c) *For every $E \in \mathcal{B}(X)$*

$$\mu(E) = \inf_{\substack{V \supseteq E \\ V \text{ open}}} \mu(V).$$

The property (c) is a consequence of (b), since for each $E \in \mathcal{B}(X)$ and $\varepsilon > 0$ there is a compact $K \subseteq X \setminus E$ such that

$$\mu(X \setminus E) < \mu(K) + \varepsilon.$$

But then

$$\mu(X \setminus K) < \mu(E) + \varepsilon$$

and $X \setminus K$ is open and contains E. In a similar way, (b) follows from (c) since X is compact.

The proof of the Riesz Representation Theorem depends on properties of continuous functions of independent interest. Suppose $K \subseteq X$ is compact and $V \subseteq X$ is open. If $f : X \to [0,1]$ is a continuous function such that

$$f \leq \chi_V \text{ and } \operatorname{supp} f \subseteq V$$

we write

$$f \prec V$$

and if

$$\chi_K \leq f \leq \chi_V \text{ and } \operatorname{supp} f \subseteq V$$

we write

$$K \prec f \prec V.$$

Theorem 3.2.2. *Let K be compact subset X.*

(a) *Suppose $K \subseteq V$ where V is open. There exists a function f on X such that*

$$K \prec f \prec V.$$

(b) *Suppose X is compact and $K \subseteq V_1 \cup ... \cup V_n$, where K is compact and $V_1, ..., V_n$ are open. There exist functions $h_1, ..., h_n$ on X such that*

$$h_i \prec V_i, \; i = 1, ..., n$$

and

$$h_1 + ... + h_n = 1 \text{ on } K.$$

PROOF. (a) Suppose $\varepsilon = \frac{1}{2}\min_K d(\cdot, V^c)$. By Corollary 3.1.2, $\varepsilon > 0$. The continuous function $f = \Pi^X_{K,\varepsilon}$ satisfies $\chi_K \le f \le \chi_{K_\varepsilon}$, that is $K \prec f \prec K_\varepsilon$. Part (a) follows if we note that the closure $(K_\varepsilon)^-$ of K_ε is contained in V.

(b) For each $x \in K$ there exists an $r_x > 0$ such that $B(x, r_x) \subseteq V_i$ for some i. Let $U_x = B(x, \frac{1}{2}r_x)$. It is important to note that $(U_x)^- \subseteq V_i$ and $(U_x)^-$ is compact since X is compact. There exist points $x_1, ..., x_m \in K$ such that $\cup_{j=1}^m U_{x_i} \supseteq K$. If $1 \le i \le n$, let F_i denote the union of those $(U_{x_j})^-$ which are contained in V_i. By Part (a), there exist continuous functions f_i such that $F_i \prec f_i \prec V_i$, $i = 1, ..., n$. Define

$$\begin{aligned} h_1 &= f_1 \\ h_2 &= (1 - f_1) f_2 \\ &\;\;.... \\ h_n &= (1 - f_1)...(1 - f_{n-1}) f_n. \end{aligned}$$

Clearly, $h_i \prec V_i$, $i = 1, ..., n$. Moreover, by induction, we get

$$h_1 + ... + h_n = 1 - (1 - f_1)...(1 - f_{n-1})(1 - f_n).$$

Since $\cup_{i=1}^n F_i \supseteq K$ we are done.

The uniqueness in Theorem 3.2.1 is simple to prove. Suppose μ_1 and μ_2 are two measures for which the theorem holds. Fix $\varepsilon > 0$ and compact $K \subseteq X$ and choose an open set V so that $\mu_2(V) \le \mu_2(K) + \varepsilon$. If $K \prec f \prec V$,

$$\mu_1(K) = \int_X \chi_K d\mu_1 \le \int_X f d\mu_1 = Tf$$

$$= \int_X f d\mu_2 \le \int_X \chi_V d\mu_2 = \mu_2(V) \le \mu_2(K) + \varepsilon.$$

Thus $\mu_1(K) \le \mu_2(K)$. If we interchange the roles of the two measures, the opposite inequality is obtained, and the uniqueness of μ follows.

To prove the existence of the measure μ in Theorem 3.2.1, define for every open V in X,

$$\mu(V) = \sup_{f \prec V} Tf.$$

Here $\mu(\phi) = 0$ since the supremum over the empty set, by convention, equals 0. Note also that $\mu(X) = T1$. Moreover, $\mu(V_1) \leq \mu(V_2)$ if V_1 and V_2 are open and $V_1 \subseteq V_2$. Now set

$$\mu(E) = \inf_{\substack{V \supseteq E \\ V \text{ open}}} \mu(V) \text{ if } E \in \mathcal{B}(X).$$

Clearly, $\mu(E_1) \leq \mu(E_2)$, if $E_1 \subseteq E_2$ and $E_1, E_2 \in \mathcal{B}(X)$. We therefore say that μ is increasing.

Lemma 3.2.1. (a) *If* $V_1, ..., V_n$ *are open,*

$$\mu(\cup_{i=1}^n V_i) \leq \Sigma_{i=1}^n \mu(V_i).$$

(b) *If* $E_1, E_2, ... \in \mathcal{B}(X)$,

$$\mu(\cup_{i=1}^\infty E_i) \leq \Sigma_{i=1}^\infty \mu(E_i).$$

(c) *If* $K_1, ..., K_n$ *are compact and pairwise disjoint,*

$$\mu(\cup_{i=1}^n K_i) = \Sigma_{i=1}^n \mu(K_i).$$

PROOF. (a) It is enough to prove (a) for $n = 2$. To this end first choose $g \prec V_1 \cup V_2$ and then $h_i \prec V_i$, $i = 1, 2$, such that $h_1 + h_2 = 1$ on supp g. Then

$$g = h_1 g + h_2 g$$

and it follows that

$$Tg = T(h_1 g) + T(h_2 g) \leq \mu(V_1) + \mu(V_2).$$

Thus

$$\mu(V_1 \cup V_2) \leq \mu(V_1) + \mu(V_2).$$

(b) Choose $\varepsilon > 0$ and for each $i \in \mathbf{N}_+$, choose an open $V_i \supseteq E_i$ such $\mu(V_i) < \mu(E_i) + 2^{-i}\varepsilon$. Set $V = \cup_{i=1}^{\infty} V_i$ and choose $f \prec V$. Since supp$f$ is compact, $f \prec V_1 \cup ... \cup V_n$ for some n. Thus, by Part (a),

$$Tf \leq \mu(V_1 \cup ... \cup V_n) \leq \Sigma_{i=1}^{n} \mu(V_i) \leq \Sigma_{i=1}^{\infty} \mu(E_i) + \varepsilon$$

and we get

$$\mu(V) \leq \Sigma_{i=1}^{\infty} \mu(E_i)$$

since $\varepsilon > 0$ is arbitrary. But $\cup_{i=1}^{\infty} E_i \subseteq V$ and it follows that

$$\mu(\cup_{i=1}^{\infty} E_i) \leq \Sigma_{i=1}^{\infty} \mu(E_i).$$

(c) It is enough to treat the special case $n = 2$. Choose $\varepsilon > 0$. Set $\rho = d(K_1, K_2)$ and $V_1 = (K_1)_{\rho/2}$ and $V_2 = (K_2)_{\rho/2}$. There is an open set $U \supseteq K_1 \cup K_2$ such that $\mu(U) < \mu(K_1 \cup K_2) + \varepsilon$ and there are functions $f_i \prec U \cap V_i$ such that $Tf_i > \mu(U \cap V_i) - \varepsilon$ for $i = 1, 2$. Now, using that μ increases

$$\mu(K_1) + \mu(K_2) \leq \mu(U \cap V_1) + \mu(U \cap V_2)$$

$$\leq Tf_1 + Tf_2 + 2\varepsilon = T(f_1 + f_2) + 2\varepsilon.$$

Since $f_1 + f_2 \prec U$,

$$\mu(K_1) + \mu(K_2) \leq \mu(U) + 2\varepsilon \leq \mu(K_1 \cup K_2) + 3\varepsilon$$

and, by letting $\varepsilon \to 0$,

$$\mu(K_1) + \mu(K_2) \leq \mu(K_1 \cup K_2).$$

The reverse inequality follows from Part (b). The lemma is proved.

Next we introduce the class

$$\mathcal{M} = \left\{ E \in \mathcal{B}(X);\ \mu(E) = \sup_{\substack{K \subseteq E \\ K \text{ compact}}} \mu(K) \right\}$$

Since μ is increasing $\mathcal{M}$ contains every compact set. Recall that a closed set in X is compact, since X is compact. Especially, note that ϕ and $X \in \mathcal{M}$.

COMPLETION OF THE PROOF OF THEOREM 3.2.1:

CLAIM 1. $\mathcal{M}$ contains every open set.

PROOF OF CLAIM 1. Let V be open and suppose $\alpha < \mu(V)$. There exists an $f \prec V$ such that $\alpha < Tf$. If U is open and $U \supseteq K =_{def} \mathrm{supp} f$, then $f \prec U$, and hence $Tf \leq \mu(U)$. But then $Tf \leq \mu(K)$. Thus $\alpha < \mu(K)$ and Claim 1 follows since K is compact and $K \subseteq V$.

CLAIM 2. Let $(E_i)_{i=1}^{\infty}$ be a disjoint denumerable collection of members of $\mathcal{M}$ and put $E = \cup_{i=1}^{\infty} E_i$. Then

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

and $E \in \mathcal{M}$.

PROOF OF CLAIM 2. Choose $\varepsilon > 0$ and for each $i \in \mathbf{N}_+$, choose a compact $K_i \subseteq E_i$ such that $\mu(K_i) > \mu(E_i) - 2^{-i}\varepsilon$. Set $H_n = K_1 \cup ... \cup K_n$. Then, by Lemma 3.2.1 (c),

$$\mu(E) \geq \mu(H_n) = \Sigma_{i=1}^{n} \mu(K_i) > \Sigma_{i=1}^{n} \mu(E_i) - \varepsilon$$

and we get

$$\mu(E) \geq \Sigma_{i=1}^{\infty} \mu(E_i).$$

Thus, by Lemma 3.2.1 (b), $\mu(E) = \Sigma_{i=1}^{\infty} \mu(E_i)$. To prove that $E \in \mathcal{M}$, let ε be as in the very first part of the proof and choose n such that

$$\mu(E) \leq \Sigma_{i=1}^{n} \mu(E_i) + \varepsilon.$$

Then

$$\mu(E) < \mu(H_n) + 2\varepsilon$$

and this shows that $E \in \mathcal{M}$.

CLAIM 3. Suppose $E \in \mathcal{M}$ and $\varepsilon > 0$. Then there exist a compact K and an open V such that $K \subseteq E \subseteq V$ and $\mu(V \setminus K) < \varepsilon$.

PROOF OF CLAIM 3. The definitions show that there exist a compact K and an open V such that

$$\mu(V) - \frac{\varepsilon}{2} < \mu(E) < \mu(K) + \frac{\varepsilon}{2}.$$

The set $V \setminus K$ is open and $V \setminus K \in \mathcal{M}$ by Claim 1. Thus Claim 2 implies that

$$\mu(K) + \mu(V \setminus K) = \mu(V) < \mu(K) + \varepsilon$$

and we get $\mu(V \setminus K) < \varepsilon$.

CLAIM 4. If $A \in \mathcal{M}$, then $X \setminus A \in \mathcal{M}$.

PROOF OF CLAIM 4. Choose $\varepsilon > 0$. Furthermore, choose compact $K \subseteq A$ and open $V \supseteq A$ such that $\mu(V \setminus K) < \varepsilon$. Then

$$X \setminus A \subseteq (V \setminus K) \cup (X \setminus V).$$

Now, by Lemma 3.2.1 (b),

$$\mu(X \setminus A) \leq \varepsilon + \mu(X \setminus V).$$

Since $X \setminus V$ is a compact subset of $X \setminus A$, we conclude that $X \setminus A \in \mathcal{M}$.

Claims 1, 2 and 4 prove that $\mathcal{M}$ is a σ-algebra which contains all Borel sets. Thus $\mathcal{M} = \mathcal{B}(X)$.

We finally prove (a). It is enough to show that

$$Tf \leq \int_X f d\mu$$

for each $f \in C(X)$. For once this is known

$$-Tf = T(-f) \leq \int_X -f d\mu \leq -\int_X f d\mu$$

and (a) follows.

Choose $\varepsilon > 0$. Set $f(X) = [a, b]$ and choose $y_0 < y_1 < ... < y_n$ such that $y_1 = a$, $y_{n-1} = b$, and $y_i - y_{i-1} < \varepsilon$. The sets

$$E_i = f^{-1}([y_{i-1}, y_i[),\ i = 1, ..., n$$

constitute a disjoint collection of Borel sets with the union X. Now, for each i, pick an open set $V_i \supseteq E_i$ such that $\mu(V_i) \leq \mu(E_i) + \frac{\varepsilon}{n}$ and $V_i \subseteq f^{-1}(]-\infty, y_i[)$. By Theorem 3.2.2 there are functions $h_i \prec V_i$, $i = 1, ..., n$, such that $\Sigma_{i=1}^n h_i = 1$ on supp$f$ and $h_i f \prec y_i h_i$ for all i. From this we get

$$\begin{aligned}
Tf = \Sigma_{i=1}^n T(h_i f) &\leq \Sigma_{i=1}^n y_i T h_i \leq \Sigma_{i=1}^n y_i \mu(V_i) \\
&\leq \Sigma_{i=1}^n y_i \mu(E_i) + \Sigma_{i=1}^n y_i \frac{\varepsilon}{n} \\
&\leq \Sigma_{i=1}^n (y_i - \varepsilon)\mu(E_i) + \varepsilon\mu(X) + (b + \varepsilon)\varepsilon \\
&\leq \Sigma_{i=1}^n \int_{E_i} f d\mu + \varepsilon\mu(X) + (b + \varepsilon)\varepsilon \\
&= \int_X f d\mu + \varepsilon\mu(X) + (b + \varepsilon)\varepsilon.
\end{aligned}$$

Since $\varepsilon > 0$ is arbitrary, we get

$$Tf \leq \int_X f d\mu.$$

This proves Theorem 3.2.1.

It is now simple to show the existence of volume measure in $\mathbf{R}^n$. For pedagogical reasons we first discuss the so called volume measure in the unit cube $Q = [0, 1]^n$ in $\mathbf{R}^n$.

The Riemann integral

$$\int_Q f(x)dx,$$

is a positive linear functional as a function of $f \in C(Q)$. Moreover, $T1 = 1$ and the Riesz Representation Theorem gives us a Borel probability measure μ in Q such that

$$\int_Q f(x)dx = \int_Q f d\mu.$$

Suppose $A \subseteq Q$ is a closed n-cell and $i \in \mathbf{N}_+$. Then

$$\mathrm{vol}(A) \leq \int_Q \Pi^Q_{A,2^{-i}}(x)dx \leq \mathrm{vol}(A_{2^{-i}})$$

and

$$\Pi^Q_{A,2^{-i}}(x) \to \chi_A(x) \text{ as } i \to \infty$$

for every $x \in \mathbf{R}^n$. Thus

$$\mu(A) = \mathrm{vol}(A).$$

The measure μ is called the volume measure in the unit cube. In the special case $n = 2$ it is called the area measure in the unit square and if $n = 1$ it is called the linear measure in the unit interval.

PROOF OF THEOREM 1.1.1. Let $\hat{\mathbf{R}}=\mathbf{R}\cup\{-\infty,\infty\}$ be the two-point compactification of $\mathbf{R}$ introduced in Example 3.1.3 and let $\hat{\mathbf{R}}^n$ denote the product of n copies of the metric space $\hat{\mathbf{R}}$. Clearly,

$$\mathcal{B}(\mathbf{R}^n) = \left\{ A \cap \mathbf{R}^n;\ A \in \mathcal{B}(\hat{\mathbf{R}}^n) \right\}.$$

Moreover, let $w : \mathbf{R}^n \to]0,\infty[$ be a continuous map such that

$$\int_{\mathbf{R}^n} w(x)dx = 1.$$

Now we define

$$Tf = \int_{\mathbf{R}^n} f(x)w(x)dx,\ f \in C(\hat{\mathbf{R}}^n).$$

Note that $T1 = 1$. The function T is a positive linear functional on $C(\hat{\mathbf{R}}^n)$ and the Riesz Representation Theorem gives us a Borel probability measure μ on $\hat{\mathbf{R}}^n$ such that

$$\int_{\mathbf{R}^n} f(x)w(x)dx = \int_{\hat{\mathbf{R}}^n} f d\mu,\ f \in C(\hat{\mathbf{R}}^n).$$

As above we get

$$\int_A w(x)dx = \mu(A)$$

for each compact n-cell in $\mathbf{R}^n$. Thus

$$\mu(\mathbf{R}^n) = \lim_{i\to\infty} \int_{[-i,i]^n} w(x)dx = 1$$

and we conclude that μ is concentrated on $\mathbf{R}^n$. Set $\mu_0(A) = \mu(A)$, $A \in \mathcal{B}(\mathbf{R}^n)$, and

$$dm_n = \frac{1}{w} d\mu_0.$$

Then, if $f \in C_c(\mathbf{R}^n)$,

$$\int_{\mathbf{R}^n} f(x)w(x)dx = \int_{\mathbf{R}^n} f d\mu_0$$

and by replacing f by f/w,

$$\int_{\mathbf{R}^n} f(x)dx = \int_{\mathbf{R}^n} f dm_n.$$

From this $m_n(A)$ =vol(A) for every compact n-cell A and it follows that m_n is the volume measure on $\mathbf{R}^n$. Theorem 1.1.1 is proved.

$\uparrow\uparrow\uparrow$

3.3 q-Adic Expansions of Numbers in the Unit Interval

To begin with in this section we will discuss so called q-adic expansions of real numbers and give some interesting consequences. As an example of an application, we construct a one-to-one real-valued Borel map f defined on a proper interval such that the range of f is a Lebesgue null set. Another example exhibits an increasing continuous function G on the unit interval with the range equal to the unit interval such that the derivative of G is equal to zero almost everywhere with respect to Lebesgue measure. In the next section we will give more applications of q-adic expansions in connection with infinite product measures.

To simplify notation let $(\Omega, P, \mathcal{F}) = ([0,1[\,, v_{1|[0,1[}, \mathcal{B}([0,1[))$. Furthermore, let $q \geq 2$ be an integer and define a function $h : \mathbf{R} \to \{0,1,2,...,q-1\}$ of period one such that

$$h(x) = k,\ \frac{k}{q} \leq x < \frac{k+1}{q},\ k = 0,...,q-1.$$

Furthermore, set for each $n \in \mathbf{N}_+$,

$$\xi_n(\omega) = h(q^{n-1}\omega),\ 0 \leq \omega < 1.$$

Then

$$P\left[\xi_n = k\right] = \frac{1}{q},\ k = 0,...,q-1.$$

Moreover, if $k_1, ..., k_n \in \{0,1,2,...,q-1\}$, it becomes obvious on drawing a figure that

$$P\left[\xi_1 = k_1, ..., \xi_{n-1} = k_{n-1}\right] = \Sigma_{i=0}^{q-1} P\left[\xi_1 = k_1, ..., \xi_{n-1} = k_{n-1}, \xi_n = i\right]$$

where each term in the sum in the right-hand side has the same value. Thus

$$P\left[\xi_1 = k_1, ..., \xi_{n-1} = k_{n-1}\right] = qP\left[\xi_1 = k_1, ..., \xi_{n-1} = k_{n-1}, \xi_n = k_n\right]$$

and

$$P\left[\xi_1 = k_1, ..., \xi_{n-1} = k_{n-1}, \xi_n = k_n\right] = P\left[\xi_1 = k_1, ..., \xi_{n-1} = k_{n-1}\right] P\left[\xi_n = k_n\right].$$

By repetition,

$$P\left[\xi_1 = k_1, ...\xi_{n-1} = k_{n-1}, \xi_n = k_n\right] = \Pi_{i=1}^{n} P\left[\xi_i = k_i\right].$$

From this we get

$$P\left[\xi_1 \in A_1, ...\xi_{n-1} \in A_{n-1}, \xi_n \in A_n\right] = \Pi_{i=1}^{n} P\left[\xi_i \in A_i\right]$$

for all $A_1, ..., A_n \subseteq \{0, 1, 2, ..., q-1\}$.

Note that each $\omega \in [0,1[$ has a so called q-adic expansion

$$\omega = \Sigma_{i=1}^{\infty} \frac{\xi_i(\omega)}{q^i}.$$

If necessary, we write $\xi_n = \xi_n^{(q)}$ to indicate q explicitly.

Let $k_0 \in \{0, 1, 2, ..., q-1\}$ be fixed and consider the event A that a number in $[0,1[$ does not have k_0 in its q-adic expansion. The probability of A equals

$$P[A] = P[\xi_i \neq k_0, \ i = 1, 2, ...] = \lim_{n\to\infty} P[\xi_i \neq k_0, \ i = 1, 2, ..., n]$$

$$= \lim_{n\to\infty} \Pi_{i=1}^{n} P[\xi_i \neq k_0] = \lim_{n\to\infty} (\frac{q-1}{q})^n = 0.$$

In particular, if

$$D_n = \left\{ \omega \in [0,1[\,; \ \xi_i^{(3)} \neq 1, \ i = 1, ..., n \right\}.$$

then, $D = \cap_{n=1}^{\infty} D_n$ is a P-zero set.

Set

$$f(\omega) = \Sigma_{i=1}^{\infty} \frac{2\xi_i^{(2)}(\omega)}{3^i}, \ 0 \leq \omega < 1.$$

We claim that f is one-to-one. If $0 \leq \omega, \omega' < 1$ and $\omega \neq \omega'$ let n be the least i such that $\xi_i^{(2)}(\omega) \neq \xi_i^{(2)}(\omega')$; we may assume that $\xi_n^{(2)}(\omega) = 0$ and $\xi_n^{(2)}(\omega') = 1$. Then

$$f(\omega') \geq \Sigma_{i=1}^{n} \frac{2\xi_i^{(2)}(\omega')}{3^i} = \Sigma_{i=1}^{n-1} \frac{2\xi_i^{(2)}(\omega')}{3^i} + \frac{2}{3^n}$$

$$= \Sigma_{i=1}^{n-1} \frac{2\xi_i^{(2)}(\omega)}{3^i} + \Sigma_{i=n+1}^{\infty} \frac{4}{3^i} > \Sigma_{i=1}^{\infty} \frac{2\xi_i^{(2)}(\omega)}{3^i} = f(\omega).$$

Thus f is one-to-one. We next prove that $f(\Omega) = D$. To this end choose $y \in D$. If $\xi_i^{(3)}(y) = 2$ for all $i \in \mathbf{N}_+$, then $y = 1$ which is a contradiction. If $k \geq 1$ is fixed and $\xi_k^{(3)}(y) = 0$ and $\xi_i^{(3)}(y) = 2$, $i \geq k+1$, then it is readily seen that $\xi_k^{(3)}(y) = 1$ which is a contradiction. Now define

$$\omega = \Sigma_{i=1}^{\infty} \frac{\frac{1}{2}\xi_i^{(3)}(y)}{2^i}$$

and we have $f(\omega) = y$.

Let $C_n = D_n^-$, $n \in \mathbf{N}_+$. The set $C = \cap_{n=1}^{\infty} C_n$, is called the Cantor set. The Cantor set is a compact Lebesgue zero set. The construction of the Cantor set may alternatively be described as follows. First $C_0 = [0,1]$. Then trisect C_0 and remove the middle interval $\left]\frac{1}{3}, \frac{2}{3}\right[$ to obtain $C_1 = C_0 \setminus \left]\frac{1}{3}, \frac{2}{3}\right[= \left[0, \frac{1}{3}\right] \cup \left[\frac{2}{3}, 1\right]$. At the second stage subdivide each of the closed intervals of C_1 into thirds and remove from each one the middle open thirds. Then $C_2 = C_1 \setminus (\left]\frac{1}{9}, \frac{2}{9}\right[\cup \left]\frac{7}{9}, \frac{8}{9}\right[)$. What is left from C_{n-1} is C_n defined above. The set $[0,1] \setminus C_n$ is the union of $2^n - 1$ intervals numbered I_k^n, $k = 1, ..., 2^n - 1$, where the interval I_k^n is situated to the left of the interval I_l^n if $k < l$.

Suppose n is fixed and let $G_n : [0,1] \to [0,1]$ be the unique monotone increasing continuous function, which satisfies $G_n(0) = 0, G_n(1) = 1, G_n(x) = k2^{-n}$ for $x \in I_k^n$ and which is affine on each interval of $C_{n.}$ It is clear that $G_n = G_{n+1}$ on each interval I_k^n, $k = 1, ..., 2^n - 1$. Moreover, $\mid G_n - G_{n+1} \mid \leq 2^{-n-1}$ and thus

$$\mid G_n - G_{n+k} \mid \leq \Sigma_{k=n}^{n+k} \mid G_k - G_{k+1} \mid \leq 2^{-n}.$$

Let $G(x) = \lim_{n \to \infty} G_n(x)$, $0 \leq x \leq 1$. The continuous and increasing function G is constant on each removed interval and it follows that $G' = 0$ a.e. with respect to linear measure in the unit interval.The function G is called the Cantor function or Cantor-Lebesgue function.

Next we introduce the following convention, which is standard in Lebesgue integration. Let $(X, \mathcal{M}, \mu)$ be a positive measure space and suppose $A \in \mathcal{M}$ and $\mu(A^c) = 0$. If two functions $g, h \in \mathcal{L}^1(\mu)$ agree on A,

$$\int_X g d\mu = \int_X h d\mu.$$

If a function $f : A \to \mathbf{R}$ is the restriction to A of a function $g \in \mathcal{L}^1(\mu)$ we define

$$\int_X f d\mu = \int_X g d\mu.$$

Now suppose $F : \mathbf{R} \to \mathbf{R}$ is a right continuous increasing function and let μ be the unique positive Borel such that

$$\mu(]a, x]) = F(x) - F(a) \text{ if } a, x \in \mathbf{R} \text{ and } a < x.$$

If $h \in L^1(\mu)$ and $E \in \mathcal{R}$, the so called Stieltjes integral

$$\int_E h(x) dF(x)$$

is by definition equal to

$$\int_E h d\mu.$$

If $a, b \in \mathbf{R}$, $a < b$, and F is continuous at the points a and b, we define

$$\int_a^b h(x) dF(x) = \int_I h d\mu$$

where I is any interval with boundary points a and b.

The reader should note that the integral

$$\int_{\mathbf{R}} h(x) dF(x)$$

in general is different from the integral

$$\int_{\mathbf{R}} h(x) F'(x) dx.$$

For example, if G is the Cantor function and G is extended so that $G(x) = 0$ for negative x and $G(x) = 1$ for x larger than 1, clearly

$$\int_{\mathbf{R}} h(x) G'(x) dx = 0$$

since $G'(x) = 0$ a.e. $[m]$. On the other hand, if we choose $h = \chi_{[0,1]}$,

$$\int_{\mathbf{R}} h(x) dG(x) = 1.$$

3.4. Product Measures

Suppose $(X, \mathcal{M})$ and $(Y, \mathcal{N})$ are two measurable spaces. If $A \in \mathcal{M}$ and $B \in \mathcal{N}$, the set $A \times B$ is called a measurable rectangle in $X \times Y$. The product σ-algebra $\mathcal{M} \otimes \mathcal{N}$ is, by definition, the σ-algebra generated by all measurable rectangles in $X \times Y$. If we introduce the projections

$$\pi_X(x, y) = x, \ (x, y) \in X \times Y$$

and

$$\pi_Y(x, y) = y, \ (x, y) \in X \times Y,$$

the product σ-algebra $\mathcal{M} \otimes \mathcal{N}$ is the least σ-algebra $\mathcal{S}$ of subsets of $X \times Y$, which makes the maps $\pi_X : (X \times Y, \mathcal{S}) \to (X, \mathcal{M})$ and $\pi_Y : (X \times Y, \mathcal{S}) \to (Y, \mathcal{N})$ measurable, that is $\mathcal{M} \otimes \mathcal{N} = \sigma(\pi_X^{-1}(\mathcal{M}) \cup \pi_Y^{-1}(\mathcal{N}))$..

Suppose $\mathcal{E}$ generates $\mathcal{M}$, where $X \in \mathcal{E}$, and $\mathcal{F}$ generates $\mathcal{N}$, where $Y \in \mathcal{F}$. We claim that the class

$$\mathcal{E} \boxtimes \mathcal{F} = \{E \times F; E \in \mathcal{E} \text{ and } F \in \mathcal{F}\}$$

generates the σ-algebra $\mathcal{M} \otimes \mathcal{N}$. First it is clear that

$$\sigma(\mathcal{E} \boxtimes \mathcal{F}) \subseteq \mathcal{M} \otimes \mathcal{N}.$$

Moreover, the class

$$\{E \in \mathcal{M}; E \times Y \in \sigma(\mathcal{E} \boxtimes \mathcal{F})\ \} = \mathcal{M} \cap \left\{E \subseteq X;\ \pi_X^{-1}(E) \in \sigma(\mathcal{E} \boxtimes \mathcal{F})\ \right\}$$

is a σ-algebra, which contains $\mathcal{E}$ and therefore equals $\mathcal{M}$. Thus $A \times Y \in \sigma(\mathcal{E} \boxtimes \mathcal{F})$ for all $A \in \mathcal{M}$ and, in a similar way, $X \times B \in \sigma(\mathcal{E} \boxtimes \mathcal{F})$ for all $B \in \mathcal{N}$ and we conclude that $A \times B = (A \times Y) \cap (X \times B) \in \sigma(\mathcal{E} \boxtimes \mathcal{F})$ for all $A \in \mathcal{M}$ and all $B \in \mathcal{N}$. This proves that

$$\mathcal{M} \otimes \mathcal{N} \subseteq \sigma(\mathcal{E} \boxtimes \mathcal{F})$$

and it follows that

$$\sigma(\mathcal{E} \boxtimes \mathcal{F}) = \mathcal{M} \otimes \mathcal{N}.$$

Thus

$$\sigma(\mathcal{E} \boxtimes \mathcal{F}) = \sigma(\mathcal{E}) \otimes \sigma(\mathcal{F}) \text{ if } X \in \mathcal{E} \text{ and } Y \in \mathcal{F}.$$

Since the σ-algebra $\mathcal{R}_n$ is generated by all open n-cells in $\mathbf{R}^n$, we conclude that

$$\mathcal{R}_{k+n} = \mathcal{R}_k \otimes \mathcal{R}_n.$$

Given $E \subseteq X \times Y$, define

$$E_x = \{y;\ (x, y) \in E\} \text{ if } x \in X$$

and

$$E^y = \{x;\ (x, y) \in E\} \text{ if } y \in Y.$$

If $f : X \times Y \to Z$ is a function and $x \in X$, $y \in Y$, let

$$f_x(y) = f(x, y), \text{ if } y \in Y$$

and

$$f^y(x) = f(x, y), \text{ if } x \in X.$$

Theorem 3.4.1 (a) *If $E \in \mathcal{M} \otimes \mathcal{N}$, then $E_x \in \mathcal{N}$* and *$E^y \in \mathcal{M}$ for every $x \in X$ and $y \in Y$.*

(b) *If $f : (X \times Y, \mathcal{M} \otimes \mathcal{N}) \to (Z, \mathcal{O})$ is measurable, then f_x is $(\mathcal{N}, \mathcal{O})$-measurable for each $x \in X$ and f^y is $(\mathcal{M}, \mathcal{O})$-measurable for each $y \in Y$.*

Proof. (a) Let

$$\mathcal{S} = \{E \in \mathcal{M} \otimes \mathcal{N}; E_x \in \mathcal{N} \text{ and } E^y \in \mathcal{M} \text{ for every } x \in X \text{ and } y \in Y\}.$$

Clearly, $X \times Y \in \mathcal{S}$. Furthermore, if $E, E_1, E_2, ... \in \mathcal{S}$, $(E^c)_x = (E_x)^c \in \mathcal{N}$ and $(\cup_{i=1}^{\infty} E_i)_x = \cup_{i=1}^{\infty} (E_i)_x \in \mathcal{N}$ for every x in X and $(E^c)^y = (E^y)^c \in \mathcal{M}$ and $(\cup_{i=1}^{\infty} E_i)^y = \cup_{i=1}^{\infty} (E_i)^y \in \mathcal{M}$ for every y in Y. It follows that $\mathcal{S}$ is a σ-algebra. Furthermore, if $A \in \mathcal{M}$ and $B \in \mathcal{N}$, $(A \times B)_x = B \in \mathcal{N}$ if $x \in A$ and $(A \times B)_x = \phi \in \mathcal{N}$ if $x \notin A$ and $(A \times B)^y = A \in \mathcal{M}$ if $y \in B$ and $(A \times B)^y = \phi \in \mathcal{M}$ if $y \notin B$. Thus $A \times B \in \mathcal{S}$ and, accordingly from this, $\mathcal{S} = \mathcal{M} \otimes \mathcal{N}$.

(b) For any set $V \in \mathcal{O}$,

$$(f^{-1}(V))_x = (f_x)^{-1}(V)$$

and

$$(f^{-1}(V))^y = (f^y)^{-1}(V).$$

Part (b) now follows from (a).

Below an $(\mathcal{M}, \mathcal{R}_{0,\infty})$-measurable or $(\mathcal{M}, \mathcal{R})$-measurable function is simply called $\mathcal{M}$-measurable.

Theorem 3.4.2. *Suppose $(X, \mathcal{M}, \mu)$ and $(Y, \mathcal{N}, \nu)$ are positive σ-finite measurable spaces and suppose $E \in \mathcal{M} \otimes \mathcal{N}$. If*

$$f(x) = \nu(E_x) \text{ and } g(y) = \mu(E^y)$$

for every $x \in X$ and $y \in Y$, then f is $\mathcal{M}$-measurable, g is $\mathcal{N}$-measurable, and

$$\int_X f d\mu = \int_Y g d\nu.$$

Proof. We first assume that $(X, \mathcal{M}, \mu)$ and $(Y, \mathcal{N}, \nu)$ are finite positive measure spaces.

Let $\mathcal{D}$ be the class of all sets $E \in \mathcal{M} \otimes \mathcal{N}$ for which the conclusion of the theorem holds. It is clear that the class $\mathcal{G}$ of all measurable rectangles in $X \times Y$ is a subset of $\mathcal{D}$ and $\mathcal{G}$ is a π-system. Furthermore, the Beppo Levi Theorem shows that $\mathcal{D}$ is a λ-system. Therefore, using Theorem 1.2.2, $\mathcal{M} \otimes \mathcal{N} = \sigma(\mathcal{G}) \subseteq \mathcal{D}$ and it follows that $\mathcal{D} = \mathcal{M} \otimes \mathcal{N}$.

In the general case, choose a denumerable disjoint collection $(X_k)_{k=1}^{\infty}$of members of $\mathcal{M}$ and a denumerable disjoint collection $(Y_n)_{n=1}^{\infty}$of members of $\mathcal{N}$ such that

$$\cup_{k=1}^{\infty} X_k = X \text{ and } \cup_{n=1}^{\infty} Y_n = Y.$$

Set

$$\mu_k = \chi_{X_k} \mu, \; k = 1, 2, ...$$

and

$$\nu_n = \chi_{Y_n} \nu, \; n = 1, 2, ... \; .$$

Then, by the Beppo Levi Theorem, the function

$$f(x) = \int_X \Sigma_{n=1}^{\infty} \chi_E(x, y) \chi_{Y_n}(y) d\nu(y)$$

$$= \Sigma_{n=1}^{\infty} \int_X \chi_E(x, y) \chi_{Y_n}(y) d\nu(y) = \Sigma_{n=1}^{\infty} \nu_n(E_x)$$

is $\mathcal{M}$-measurable. Again, by the Beppo Levi Theorem,

$$\int_X f d\mu = \Sigma_{k=1}^{\infty} \int_X f d\mu_k$$

and

$$\int_X f d\mu = \Sigma_{k=1}^{\infty} (\Sigma_{n=1}^{\infty} \int_X \nu_n(E_x) d\mu_k(x)) = \Sigma_{k,n=1}^{\infty} \int_X \nu_n(E_x) d\mu_k(x).$$

In a similar way, the function g is $\mathcal{N}$-measurable and

$$\int_Y g d\nu = \Sigma_{n=1}^{\infty} (\Sigma_{k=1}^{\infty} \int_Y \mu_k(E^y) d\nu_n(y)) = \Sigma_{k,n=1}^{\infty} \int_Y \mu_k(E^y) d\nu_n(y).$$

Since the theorem is true for finite positive measure spaces, the general case follows.

Definition 3.4.1. If $(X, \mathcal{M}, \mu)$ and $(Y, \mathcal{N}, \nu)$ are positive σ-finite measurable spaces and $E \in \mathcal{M} \otimes \mathcal{N}$, define

$$(\mu \times \nu)(E) = \int_X \nu(E_x) d\mu(x) = \int_Y \mu(E^y) d\nu(y).$$

The function $\mu \times \nu$ is called the product of the measures μ and ν.

Note that Beppo Levi's Theorem ensures that $\mu \otimes \nu$ is a positive measure.

Before the next theorem we recall the following convention. Let $(X, \mathcal{M}, \mu)$ be a positive measure space and suppose $A \in \mathcal{M}$ and $\mu(A^c) = 0$. If two functions $g, h \in \mathcal{L}^1(\mu)$ agree on A,

$$\int_X g d\mu = \int_X h d\mu.$$

If a function $f : A \to \mathbf{R}$ is the restriction to A of a function $g \in \mathcal{L}^1(\mu)$ we define

$$\int_X f d\mu = \int_X g d\mu.$$

Theorem 3.4.3. *Let $(X, \mathcal{M}, \mu)$ and $(Y, \mathcal{N}, \nu)$ be positive σ-finite measurable spaces.*

(a) **(Tonelli's Theorem)** *If $h : X \times Y \to [0, \infty]$ is $(\mathcal{M} \otimes \mathcal{N})$-measurable and*

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

for every $x \in X$ and $y \in Y$, then f is $\mathcal{M}$-measurable, g is $\mathcal{N}$-measurable, and

$$\int_X f d\mu = \int_{X \times Y} h d(\mu \times \nu) = \int_Y g d\nu$$

(b) **(Fubini's Theorem)**

(i) If $h : X \times Y \to \mathbf{R}$ *is* $(\mathcal{M} \otimes \mathcal{N})$*-measurable and*

$$\int_X (\int_Y \mid h(x,y) \mid d\nu(y)) d\mu(x) < \infty$$

then $h \in L^1(\mu \times \nu)$*. Moreover,*

$$\int_X (\int_Y h(x,y) d\nu(y)) d\mu(x) = \int_{X \times Y} h d(\mu \times \nu) = \int_Y (\int_X h(x,y) d\mu(x)) d\nu(y)$$

(ii) If $h \in L^1((\mu \times \nu)^-)$*, then* $h_x \in L^1(\nu)$ *for* μ*-almost all* x *and*

$$\int_{X \times Y} h d(\mu \times \nu) = \int_X (\int_Y h(x,y) d\nu(y)) d\mu(x)$$

(iii) If $h \in L^1((\mu \times \nu)^-)$*, then* $h^y \in L^1(\mu)$ *for* ν*-almost all* y *and*

$$\int_{X \times Y} h d(\mu \times \nu) = \int_Y (\int_X h(x,y) d\mu(x)) d\nu(y)$$

PROOF. (a) The special case when h is a non-negative $(\mathcal{M} \otimes \mathcal{N})$-measurable simple function follows from Theorem 3.4.2. Remembering that any non-negative measurable function is the pointwise limit of an increasing sequence of simple measurable functions, the Lebesgue Monotone Convergence Theorem implies the Tonelli Theorem.

(b) PART (i) : By Part (a)

$$\infty > \int_X (\int_Y h^+(x,y) d\nu(y)) d\mu(x) = \int_{X \times Y} h^+ d(\mu \times \nu)$$

$$= \int_Y (\int_X h^+(x,y) d\mu(x)) d\nu(y)$$

and

$$\infty > \int_X (\int_Y h^-(x,y) d\nu(y)) d\mu(x) = \int_{X \times Y} h^- d(\mu \times \nu)$$

$$= \int_Y (\int_X h^-(x,y) d\mu(x)) d\nu(y).$$

Let

$$A = \left\{ x \in X;\ (h^+)_x, (h^-)_x \in L^1(\nu) \right\}.$$

Then A^c is a μ-null set and we get

$$\int_A (\int_Y h^+(x,y)d\nu(y))d\mu(x) = \int_{X\times Y} h^+ d(\mu\times\nu)$$

and

$$\int_A (\int_Y h^-(x,y)d\nu(y))d\mu(x) = \int_{X\times Y} h^- d(\mu\times\nu).$$

Thus

$$\int_A (\int_Y h(x,y)d\nu(y))d\mu(x) = \int_{X\times Y} hd(\mu\times\nu)$$

and, hence,

$$\int_X (\int_Y h(x,y)d\nu(y))d\mu(x) = \int_{X\times Y} hd(\mu\times\nu).$$

The other case can be treated in a similar way. The theorem is proved.

PART (ii) : We first use Theorem 2.2.3 and write $h = \varphi + \psi$ where $\varphi \in L^1(\mu\times\nu)$, ψ is $(\mathcal{M}\otimes\mathcal{N})^-$-measurable and $\psi = 0$ a.e. $[\mu\times\nu]$. Set

$$A = \left\{ x \in X;\ (\varphi^+)_x, (\varphi^-)_x \in L^1(\nu) \right\}.$$

Furthermore, suppose $E \supseteq \{(x,y);\ \psi(x,y) \neq 0\}$, $E \in \mathcal{M}\otimes\mathcal{N}$ and

$$(\mu\times\nu)(E) = 0.$$

Then, by Tonelli's Theorem

$$0 = \int_X \nu(E_x)d\mu(x).$$

Let $B = \{x \in X;\ \nu(E_x) \neq 0\}$ and note that $B \in \mathcal{M}$. Moreover $\mu(B) = 0$ and if $x \notin B$, then $\psi_x = 0$ a.e. $[\nu]$ that is $h_x = \varphi_x$ a.e. $[\nu]$. Now, by Part (i)

$$\int_{X\times Y} hd(\mu\times\nu)^- = \int_{X\times Y} \varphi d(\mu\times\nu) = \int_A (\int_Y \varphi(x,y)d\nu(y))d\mu(x)$$

$$= \int_{A\cap B^c} (\int_Y \varphi(x,y)d\nu(y))d\mu(x) = \int_{A\cap B^c} (\int_Y h(x,y)d\nu(y))d\mu(x)$$
$$= \int_X (\int_Y h(x,y)d\nu(y))d\mu(x).$$

Part (iii) is proved in the same manner as Part (ii). This concludes the proof of the theorem.

If $(X_i, \mathcal{M}_i)$, $i = 1, ..., n$, are measurable spaces, the product σ-algebra $\mathcal{M}_1 \otimes ... \otimes \mathcal{M}_n$ is, by definition, the σ-algebra generated by all sets of the form

$$A_1 \times ... \times A_n$$

where $A_i \in \mathcal{M}_i$, $i = 1, ..., n$. Now assume $(X_i, \mathcal{M}_i, \mu_i)$, $i = 1, ..., n$, are σ-finite positive measure spaces. By induction, we define $\nu_1 = \mu_1$ and $\nu_k = \nu_{k-1} \times \mu_k$, $k = 1, 2, ..., n$. The measure, ν_n is called the product of the measures $\mu_1, ..., \mu_n$ and is denoted by $\mu_1 \times ... \times \mu_n$. It is readily seen that

$$\mathcal{R}_n = \mathcal{R}_1 \otimes ... \otimes \mathcal{R}_1 \text{ (n factors)}$$

and

$$v_n = v_1 \times ... \times v_1 \text{ (n factors).}$$

Moreover,

$$\mathcal{R}_n^- \supseteq (\mathcal{R}_1^-)^n =_{def} \mathcal{R}_1^- \otimes ... \otimes \mathcal{R}_1^- \text{ (n factors).}$$

If $A \in \mathcal{P}(\mathbf{R}) \setminus \mathcal{R}_1^-$, by the Tonelli Theorem, the set $A \times \{0, ..., 0\}$ ($n-1$ zeros) is an m_n-null set, which, in view of Theorem 3.4.1, cannot belong to the σ-algebra $(\mathcal{R}_1^-)^n$. Thus the Axiom of Choice implies that

$$\mathcal{R}_n^- \neq (\mathcal{R}_1^-)^n.$$

Clearly, the completion of the measure $m_1 \times ... \times m_1$ (n factors) equals m_n.

Sometimes we prefer to write

$$\int_{A_1 \times ... \times A_n} f(x_1, ..., x_n)dx_1...dx_n$$

instead of

$$\int_{A_1 \times ... \times A_n} f(x) dm_n(x)$$

or

$$\int_{A_1 \times ... \times A_n} f(x) dx.$$

Moreover, the integral

$$\int_{A_1} ... \int_{A_n} f(x_1, ..., x_n) dx_1 ... dx_n$$

is the same as

$$\int_{A_1 \times ... \times A_n} f(x_1, ..., x_n) dx_1 ... dx_n.$$

Definition 3.4.2. (a) The measure

$$\gamma_1(A) = \int_A e^{-\frac{x^2}{2}} \frac{dx}{\sqrt{2\pi}}, \ A \in \mathcal{R}$$

is called the canonical Gauss measure in $\mathbf{R}$.

(b) The measure

$$\gamma_n = \gamma_1 \times ... \times \gamma_1 \ (n \text{ factors})$$

is called the canonical Gauss measure in $\mathbf{R}^n$. Thus, if

$$\mid x \mid = \sqrt{x_1^2 + ... + x_n^2}, \ x = (x_1, ..., x_n) \in \mathbf{R}^n$$

we have

$$\gamma_n(A) = \int_A e^{-\frac{|x|^2}{2}} \frac{dx}{\sqrt{2\pi}}, \ A \in \mathcal{R}_n.$$

(c) A Borel measure μ in $\mathbf{R}$ is said to be a centred Gaussian measure if $\mu = f(\gamma_1)$ for some linear map $f : \mathbf{R} \to \mathbf{R}$.

(d) A real-valued random variable ξ is said to be a centred Gaussian random variable if its probability law is a centred Gaussian measure in $\mathbf{R}$. Stated otherwise, ξ is a real-valued centred Gaussian random variable if either

$$\mathcal{L}(\xi) = \delta_0 \text{ (abbreviated } \xi \in N(0,0))$$

or there exists a $\sigma > 0$ such that

$$\mathcal{L}(\frac{\xi}{\sigma}) = \gamma_1 \text{ (abbreviated } \xi \in N(0,\sigma)).$$

(e) A family $(\xi_t)_{t\in T}$ of real-valued random variables is said to be a centred real-valued Gaussian process if for all $t_1, ..., t_n \in T$, $\alpha_1, ..., \alpha_n \in \mathbf{R}$ and every $n \in \mathbf{N}_+$, the sum

$$\xi = \Sigma_{k=1}^n \alpha_k \xi_{t_k}$$

is a centred Gaussian random variable.

Exercises

1. Let $(X, \mathcal{M}, \mu)$ and $(Y, \mathcal{N}, \nu)$ be two σ-finite measure spaces. Let $f \in L^1(\mu)$ and $g \in L^1(\nu)$ and define $h(x,y) = f(x)g(y)$, $(x,y) \in X \times Y$. Prove that $h \in L^1(\mu \times \nu)$ and

$$\int_{X\times Y} hd(\mu\times\nu) = \int_X fd\mu \int_Y gd\nu.$$

2. Let $(X, \mathcal{M}, \mu)$ be a σ-finite measure space and $f : X \to [0, \infty[$ a measurable function. Prove that

$$\int_X fd\mu = (\mu \times m)\left\{(x,y);\ 0 < y < f(x),\ x \in X\right\}.$$

3. Let $(X, \mathcal{M}, \mu)$ be a σ-finite measure space and $f : X \to \mathbf{R}$ a measurable function. Prove that $(\mu \times m)(\{(x, f(x));\ x \in X\ \}) = 0$.

4. Let $E \in \mathcal{R}_2^-$ and $E \subseteq [0,1] \times [0,1]$. Suppose $m(E_x) \leq \frac{1}{2}$ for m-almost all $x \in [0,1]$. Show that

$$m(\{y \in [0,1]\,; m(E^y) = 1\}) \leq \frac{1}{2}.$$

5. Let c be the counting measure on $\mathbf{R}$ restricted to $\mathcal{R}$ and

$$D = \{(x,x);\ x \in \mathbf{R}\}\,.$$

Define for every $A \in (\mathcal{R} \boxtimes \mathcal{R}) \cup \{D\}$,

$$\mu(A) = \int_{\mathbf{R}} (\int_{\mathbf{R}} \chi_A(x,y) dv_1(x)) dc(y)$$

and

$$\nu(A) = \int_{\mathbf{R}} (\int_{\mathbf{R}} \chi_A(x,y) dc(y)) dv_1(x).$$

(a) Prove that μ and ν agree on $\mathcal{R} \boxtimes \mathcal{R}$.
(b) Prove that $\mu(D) \neq \nu(D)$.

6. Let $I =]0,1[$ and

$$h(x,y) = \frac{x^2 - y^2}{(x^2+y^2)^2},\ (x,y) \in I \times I.$$

Prove that

$$\int_I (\int_I h(x,y) dy) dx = \frac{\pi}{4},$$

$$\int_I (\int_I h(x,y) dx) dy = -\frac{\pi}{4}$$

and

$$\int_{I \times I} \mid h(x,y) \mid dx dy = \infty.$$

7. For $t > 0$ and $x \in \mathbf{R}$ let

$$g(t,x) = \frac{1}{\sqrt{2\pi t}} e^{-\frac{x^2}{2t}}$$

and

$$h(t,x) = \frac{\partial g}{\partial t}.$$

Given $a > 0$, prove that

$$\int_{-\infty}^{\infty} (\int_{a}^{\infty} h(t,x)dt)dx = -1$$

and

$$\int_{a}^{\infty} (\int_{-\infty}^{\infty} h(t,x)dx)dt = 0$$

and conclude that

$$\int_{[a,\infty[\times\mathbf{R}} \mid h(t,x) \mid dtdx = \infty.$$

(Hint: First prove that

$$\int_{-\infty}^{\infty} g(t,x)dx = 1$$

and

$$\frac{\partial g}{\partial t} = \frac{1}{2}\frac{\partial^2 g}{\partial x^2}.)$$

8. Given $f \in L^1(m)$, let

$$g(x) = \frac{1}{2}\int_{x-1}^{x+1} f(t)dt,\ x \in \mathbf{R}.$$

Prove that

$$\int_{\mathbf{R}} \mid g(x) \mid dx \leq \int_{\mathbf{R}} \mid f(x) \mid dx.$$

9. Let $I = [0,1]$ and suppose $f : I \to \mathbf{R}$ is a Lebesgue measurable function such that

$$\int_{I\times I} \mid f(x) - f(y) \mid dxdy < \infty.$$

Prove that

$$\int_{I} \mid f(x) \mid dx < \infty.$$

10. Suppose $A \in \mathcal{R}^-$ and $f \in L^1(m)$. Set

$$g(x) = \int_{\mathbf{R}} \frac{d(y,A)f(y)}{\mid x-y \mid^2}dy,\ x \in \mathbf{R}.$$

Prove that

$$\int_A | g(x) | \, dx < \infty.$$

11. Suppose that the functions $f, g : \mathbf{R} \to [0, \infty[$ are Lebesgue measurable and introduce $\mu = fm$ and $\nu = gm$. Prove that the measures μ and ν are σ-finite and

$$(\mu \times \nu)(E) = \int_E f(x)g(y)dxdy \text{ if } E \in \mathcal{R}^- \otimes \mathcal{R}^-.$$

12. Suppose μ is a finite positive Borel measure on $\mathbf{R}^n$ and $f : \mathbf{R}^n \to \mathbf{R}$ a Borel measurable function. Set $g(x, y) = f(x) - f(y)$, $x, y \in \mathbf{R}^n$. Prove that $f \in L^1(\mu)$ if and only if $g \in L^1(\mu \times \mu)$.

13. A random variable ξ is non-negative and possesses the distribution function $F(x) = P[\xi \leq x]$. Prove that$E[\xi] = \int_0^\infty (1 - F(x))dx$.

14. Let (X, d) be a metric space and suppose $Y \in \mathcal{B}(X)$. Then Y equipped with the metric $d_{|Y \times Y}$ is a metric space. Prove that

$$\mathcal{B}(Y) = \{A \cap Y; \ A \in \mathcal{B}(X)\}.$$

15. The continuous bijection $f : (X, d) \to (Y, e)$ has a continuous inverse. Prove that $f(A) \in \mathcal{B}(Y)$ if $A \in \mathcal{B}(X)$

16. A real-valued function $f(x, y), x, y \in \mathbf{R}$, is a Borel function of x for every fixed y and a continuous function of y for every fixed x. Prove that f is a Borel function. Is the same conclusion true if we only assume that $f(x, y)$ is a real-valued Borel function in each variable separately?

17. Suppose $a > 0$ and

$$\mu_a = e^{-a} \sum_{n=0}^{\infty} \frac{a^n}{n!} \delta_n$$

where $\delta_n(A) = \chi_A(n)$ if $n \in \mathbf{N} = \{0, 1, 2, ...\}$ and $A \subseteq \mathbf{N}$. Prove that

$$(\mu_a \times \mu_b) s^{-1} = \mu_{a+b}$$

for all $a, b > 0$, if $s(x, y) = x + y$, $x, y \in \mathbf{N}$.

18. Suppose

$$f(t) = \int_0^{\infty} \frac{x e^{-x^2}}{x^2 + t^2} dx, \ t > 0.$$

Compute

$$\lim_{t \to 0+} f(t) \text{ and } \int_0^{\infty} f(t) dt.$$

Finally, prove that f is differentiable.

3.5 Change of Variables in Volume Integrals

If T is a non-singular n by n matrix with real entries, we claim that

$$T(v_n) = \frac{1}{\mid \det T \mid} v_n$$

(here T is viewed as a linear map of $\mathbf{R}^n$ into $\mathbf{R}^n$). Remembering Corollary 3.1.3 this means that the following linear change of variables formula holds, viz.

$$\int_{\mathbf{R}^n} f(Tx) dx = \frac{1}{\mid \det T \mid} \int_{\mathbf{R}^n} f(x) dx \text{ all } f \in C_c(\mathbf{R}^n).$$

The case $n = 1$ is obvious. Moreover, by Fubini's Theorem the linear change of variables formula is true for arbitrary n in the following cases:

(a) $Tx = (x_{\pi(1)}, ..., x_{\pi(n)})$, where π is a permutation of the numbers $1, ..., n$.

(b) $Tx = (\alpha x_1, x_2, ..., x_n)$, where α is a non-zero real number.

(c) $Tx = (x_1 + x_2, x_2, ..., x_n)$.

Recall from linear algebra that every non-singular n by n matrix T can be row-reduced to the identity matrix, that is T can by written as the product of finitely many transformations of the types in (a),(b), and (c). This proves the above linear change of variables formula.

Our main objective in this section is to prove a more general change of variable formula. To this end let Ω and Γ be open subsets of $\mathbf{R}^n$ and $G : \Omega \to \Gamma$ a C^1 diffeomorphism, that is $G = (g_1, ..., g_n)$ is a bijective continuously differentiable map such that the matrix $G'(x) = (\frac{\partial g_i}{\partial x_j}(x))_{1 \leq i,j \leq n}$ is non-singular for each $x \in \Omega$. The inverse function theorem implies that $G^{-1} : \Gamma \to \Omega$ is a C^1 diffeomorphism $[DI]$.

Theorem 3.5.1. *If f is a non-negative Borel function in Ω, then*

$$\int_\Gamma f(x)dx = \int_\Omega f(G(x)) \mid \det G'(x) \mid dx.$$

The proof of Theorem 3.5.1 is based on several lemmas.

Throughout, $\mathbf{R}^n$ is equipped with the metric

$$d_n(x, y) = \max_{1 \leq k \leq n} \mid x_k - y_k \mid .$$

Let K be a compact convex subset of Ω. Then if $x, y \in K$ and $1 \leq i \leq n$,

$$g_i(x) - g_i(y) = \int_0^1 \frac{d}{dt} g_i(y + t(x - y))dt$$

$$= \int_0^1 \Sigma_{k=1}^n \frac{\partial g_i}{\partial x_k}(y + t(x - y))(x_k - y_k)dt$$

and we get

$$d_n(G(x), G(y)) \leq M(G, K)d_n(x, y)$$

where

$$M(G, K) = \max_{1 \leq i \leq n} \Sigma_{k=1}^n \max_{z \in K} \mid \frac{\partial g_i}{\partial x_k}(x) \mid .$$

Thus if $\bar{B}(a; r)$ is a closed ball contained in K,

$$G(\bar{B}(a; r)) \subseteq \bar{B}(G(a); M(G, K)r).$$

Lemma 3.5.1. *Let $(Q_k)_{k=1}^{\infty}$ be a sequence of closed balls contained in Ω such that*

$$Q_{k+1} \subseteq Q_k$$

and

$$\operatorname{diam} Q_k \to 0 \text{ as } k \to \infty.$$

Then, there is a unique point a belonging to each Q_k and

$$\limsup_{n\to\infty} \frac{v_n(G(Q_k))}{v_n(Q_k)} \leq \mid \det G'(a) \mid .$$

PROOF. The existence of a point a belonging to each Q_k is an immediate consequence of Theorem 3.1.2. The uniqueness is also obvious since the diameter of Q_k converges to 0 as $k \to \infty$. Set $T = G'(a)$ and $F = T^{-1}G$. Then, if $Q_k = \bar{B}(x_k; r_k)$,

$$v_n(G(Q_k)) = v_n(T(T^{-1}G(Q_k))) = \mid \det T \mid \ v_n(T^{-1}G(\bar{B}(x_k; r_k)))$$

$$\leq \mid \det T \mid v_n(\bar{B}(T^{-1}G(x_k); M(T^{-1}G; Q_k)r_k) = \mid \det T \mid \ M(T^{-1}G; Q_k)^n v_n(Q_k).$$

Since

$$\lim_{k\to\infty} M(T^{-1}G; Q_k) = 1$$

the lemma follows at once.

Lemma 3.5.2. *Let Q be a closed ball contained in Ω. Then*

$$v_n(G(Q)) \leq \int_Q \mid \det G'(x) \mid dx.$$

PROOF. Suppose there is a closed ball Q contained in Ω such that

$$v_n(G(Q)) > \int_Q \mid \det G'(x) \mid dx.$$

This will lead us to a contradiction as follows.

Choose $\varepsilon > 0$ such that

$$v_n(G(Q)) \geq (1+\varepsilon) \int_Q \mid \det G'(x) \mid dx.$$

Let $Q = \cup_1^{2^n} B_k$ where $B_1, ..., B_{2^n}$ are mutually almost disjoint closed balls with the same volume. If

$$v_n(G(B_k)) < (1+\varepsilon)\int_{B_k} \mid \det G'(x) \mid dx,\ \ k = 1, ..., 2^n$$

we get

$$v_n(G(Q)) \leq \Sigma_{k=1}^{2^n} v_n(G(B_k))$$

$$< \Sigma_{k=1}^{2^n}(1+\varepsilon)\int_{B_k} \mid \det G'(x) \mid dx = (1+\varepsilon)\int_Q \mid \det G'(x) \mid dx$$

which is a contradiction. Thus

$$v_n(G(B_k)) \geq (1+\varepsilon)\int_{B_k} \mid \det G'(x) \mid dx$$

for some k. By induction we obtain a sequence $(Q_k)_{k=1}^{\infty}$ of closed balls contained in Ω such that

$$Q_{k+1} \subseteq Q_k,$$

$$\text{diam } Q_k \to 0 \text{ as } k \to \infty$$

and

$$v_n(G(Q_k)) \geq (1+\varepsilon)\int_{Q_k} \mid \det G'(x) \mid dx.$$

But applying Lemma 3.5.1 we get a contradiction.

PROOF OF THEOREM 3.5.1. Let $U \subseteq \Omega$ be open and write $U = \cup_{i=1}^{\infty} Q_i$ where the $Q_i's$ are almost disjoint cubes as in Theorem 3.1.5. Then

$$v_n(G(U)) \leq \Sigma_{i=1}^{\infty} v_n(G(Q_i)) \leq \Sigma_{i=1}^{\infty}\int_{Q_i} \mid \det G'(x) \mid dx$$

$$= \int_U \mid \det G'(x) \mid dx.$$

Using Theorem 3.1.3 we now have that

$$v_n(G(E)) \leq \int_E \mid \det G'(x) \mid dx$$

for each Borel set $E \subseteq \Omega$. But then

$$\int_\Gamma f(x)dx \leq \int_\Omega f(G(x)) \mid \det G'(x) \mid dx$$

for each simple Borel measurable function $f \geq 0$ and, accordingly from this and monotone convergence, the same inequality holds for each non-negative Borel function f. But the same line of reasoning applies to G replaced by G^{-1} and f replaced by $f(G) \mid \det G' \mid$, so that

$$\int_\Omega f(G(x)) \mid \det G'(x) \mid dx \leq \int_\Gamma f(x) \mid \det G'(G^{-1}(x)) \mid\mid \det(G^{-1})'(x) \mid dx$$

$$= \int_\Gamma f(x)dx.$$

This proves the theorem.

Example 3.5.1. If $f : \mathbf{R}^2 \to [0, \infty]$ is $(\mathcal{R}_2, \mathcal{R}_{0,\infty})$-measurable and $0 < \varepsilon < R < \infty$, the substitution

$$G(r, \theta) = (r \cos\theta, r \sin\theta)$$

yields

$$\int_{\varepsilon < \sqrt{x_1^2 + x_2^2} < R} f(x_1, x_2)dx_1dx_2 = \int_\varepsilon^R \int_0^{2\pi} f(r\cos\theta, r\sin\theta)rdrd\theta$$

and by letting $\varepsilon \to 0$ and $R \to \infty$, we have

$$\int_{\mathbf{R}^2} f(x_1, x_2)dx_1dx_2 = \int_0^\infty \int_0^{2\pi} f(r\cos\theta, r\sin\theta)rdrd\theta.$$

The purpose of the example is to show an analogue formula for volume measure in $\mathbf{R}^n$.

Let $S^{n-1} = \{x \in \mathbf{R}^n; \mid x \mid= 1\}$ be the unit sphere in $\mathbf{R}^n$. We will define a so called surface area Borel measure σ_{n-1} on S^{n-1} such that

$$\int_{\mathbf{R}^n} f(x)dx = \int_0^\infty \int_{S^{n-1}} f(r\omega)r^{n-1}drd\sigma_{n-1}(\omega)$$

for any $(\mathcal{R}_n, \mathcal{R}_{0,\infty})$-measurable function $f : \mathbf{R}^n \to [0, \infty]$. To this end define $G : \mathbf{R}^n \setminus \{0\} \to]0, \infty[\times S^{n-1}$ by setting $G(x) = (r, \omega)$, where

$$r =\mid x \mid \text{ and } \omega = \frac{x}{\mid x \mid}.$$

Note that $G^{-1} :]0, \infty[\times S^{n-1} \to \mathbf{R}^n \setminus \{0\}$ is given by the equation

$$G^{-1}(r, \omega) = r\omega.$$

Moreover,

$$G^{-1}(]0, a] \times E) = aG^{-1}(]0, 1] \times E) \text{ if } a > 0 \text{ and } E \subseteq S^{n-1}.$$

If $E \in \mathcal{B}(S^{n-1})$ we therefore have that

$$v_n(G^{-1}(]0, a] \times E)) = a^n v_n(G^{-1}(]0, 1] \times E)).$$

We now define

$$\sigma_{n-1}(E) = nv_n(G^{-1}(]0, 1] \times E)) \text{ if } E \in \mathcal{B}(S^{n-1})$$

and

$$\rho(A) = \int_A r^{n-1} dr \text{ if } A \in \mathcal{B}(]0, \infty[).$$

Below, by abuse of language, we write $v_{n|\mathbf{R}^n \setminus \{0\}} = v_n$. Then, if $0 < a \leq b < \infty$ and $E \in \mathcal{B}(S^{n-1})$,

$$G(v_n)(]0, a] \times E) = \rho(]0, a])\sigma_{n-1}(E)$$

and

$$G(v_n)(]a, b] \times E) = \rho(]a, b])\sigma_{n-1}(E).$$

Thus, by Theorem 1.2.3,

$$G(v_n) = \rho \times \sigma_{n-1}$$

and the claim above is immediate.

To check the normalization constant in the definition of σ_{n-1}, first note that

$$v_n(\mid x \mid < R) = \int_0^R \int_{S^{n-1}} r^{n-1} dr d\sigma(\omega) = \frac{R^n}{n} \sigma_{n-1}(S^{n-1})$$

and we get

$$\frac{d}{dR}v_n(\mid x \mid < R) = R^{n-1}\sigma_{n-1}(S^{n-1}).$$

Exercises

1. Extend Theorem 3.5.1 to Lebesgue measurable functions.

2. The function $f : \mathbf{R} \to [0, \infty[$ is Lebesgue measurable and $\int_{\mathbf{R}} f dm = 1$. Determine all non-zero real numbers α such that $\int_{\mathbf{R}} h dm < \infty$, where

$$h(x) = \Sigma_{n=0}^{\infty} f(\alpha^n x + n),\ \ x \in \mathbf{R}.$$

$\downarrow\downarrow\downarrow$

3.6. Independence in Probability

Suppose $(\Omega, \mathcal{F}, P)$ is a probability space. The random variables $\xi_k : (\Omega, P) \to (S_k, \mathcal{S}_k),\ k = 1, ..., n$ are said to be independent if

$$P_{(\xi_1,...,\xi_n)} = \times_{k=1}^{n} P_{\xi_k}.$$

A family $(\xi_i)_{i\in I}$ of random variables is said to be independent if $\xi_{i_1}, ..., \xi_{i_n}$ are independent for any $i_1, ...i_n \in I$ with $i_k \neq i_l$ if $k \neq l$. A family of events $(A_i)_{i\in I}$ is said to be independent if $(\chi_{A_i})_{i\in I}$ is a family of independent random variables. Finally a family $(\mathcal{A}_i)_{i\in I}$ of sub-σ-algebras of $\mathcal{F}$ is said to be independent if, for any $A_i \in \mathcal{A}_i,\ i \in I$, the family $(A_i)_{i\in I}$ is a family of independent events.

Example 3.6.1. Let $q \geq 2$ be an integer. A real number $\omega \in [0, 1[$ has a q-adic expansion

$$\omega = \Sigma_{k=1}^{\infty} \frac{\xi_k^{(q)}}{q^k}.$$

The construction of the Cantor set shows that $(\xi_k^{(q)})_{k=1}^{\infty}$ is a sequence of independent random variables based on the probability space

$$([0,1[\,, v_{1|[0,1[}, \mathcal{B}([0,1[)).$$

Example 3.6.2. Let $(X, \mathcal{M}, \mu)$ be a positive measure space and let $A_i \in \mathcal{M}$, $i \in \mathbf{N}_+$, be such that

$$\Sigma_{i=1}^{\infty} \mu(A_i) < \infty.$$

The first Borel-Cantelli Lemma asserts that μ-almost all $x \in X$ lie in A_i for at most finitely many i. This result is an immediate consequence of the Beppo Levi Theorem since

$$\int_X \Sigma_{i=1}^{\infty} \chi_{A_i} d\mu = \Sigma_{i=1}^{\infty} \int_X \chi_{A_i} d\mu < \infty$$

implies that

$$\Sigma_{i=1}^{\infty} \chi_{A_i} < \infty \text{ a.e. } [\mu]\,.$$

Suppose $(\Omega, \mathcal{F}, P)$ is a probability space and let $(A_i)_{i=1}^{\infty}$ be independent events such that

$$\Sigma_{i=1}^{\infty} P[A_i] = \infty.$$

The second Borel-Cantelli Lemma asserts that almost surely A_i happens for infinitely many i.

To prove this, we use the inequality

$$1 + x \leq e^x,\ x \in \mathbf{R}$$

to obtain

$$P\left[\cap_{i=k}^{k+n} A_i^c\right] = \Pi_{i=k}^{k+n} P\left[A_i^c\right]$$

$$= \Pi_{i=k}^{k+n}(1 - P[A_i]) \leq \Pi_{i=k}^{k+n} e^{-P[A_i]} = e^{-\Sigma_{i=k}^{k+n} P[A_i]}.$$

By letting $n \to \infty$,

$$P\left[\cap_{i=k}^{\infty} A_i^c\right] = 0$$

or

$$P\left[\cup_{i=k}^{\infty} A_i\right] = 1.$$

But then

$$P\left[\cap_{k=1}^{\infty} \cup_{i=k}^{\infty} A_i\right] = 1$$

and the second Borel-Cantelli Lemma is proved.

Theorem 3.6.1. *Suppose* $\xi_1, ..., \xi_n$ *are independent random variables and* $\xi_k \in N(0,1),\ k = 1, ..., n$. *If* $\alpha_1, ..., \alpha_n \in R$, *then*

$$\Sigma_{k=1}^{n} \alpha_k \xi_k \in N(0, \Sigma_{k=1}^{n} \alpha_k^2)$$

PROOF. The case $\alpha_1, ..., \alpha_n = 0$ is trivial so assume $\alpha_k \neq 0$ for some k. We have for each open interval A,

$$P\left[\Sigma_{k=1}^{n} \alpha_k \xi_k \in A\right] = \int_{\Sigma_{k=1}^{n} \alpha_k x_k \in A} d\gamma_1(x_1)...d\gamma_1(x_n)$$

$$\int_{\Sigma_{k=1}^{n} \alpha_k x_k \in A} \frac{1}{\sqrt{2\pi}^n} e^{-\frac{1}{2}(x_1^2 + ... + x_n^2)} dx_1...dx_n.$$

Set $\sigma = \sqrt{\alpha_1^2 + ... + \alpha_n^2}$ and let $y = Gx$ be an orthogonal transformation such that

$$y_1 = \frac{1}{\sigma}(\alpha_1 x_1 + ... + \alpha_n x_n).$$

Then, since $\det G = 1$,

$$P\left[\Sigma_{k=1}^{\infty} \alpha_k \xi_k \in A\right] = \int_{\sigma y_1 \in A} \frac{1}{\sqrt{2\pi}^n} e^{-\frac{1}{2}(y_1^2 + ... + y_n^2)} dy_1...dy_n$$

$$= \int_{\sigma y_1 \in A} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}y_1^2} dy_1$$

where we used Fubini's theorem in the last step. The theorem is proved.

Finally, in this section, we prove a basic result about the existence of infinite product measures. Let $\mu_k,\ k \in \mathbf{N}_+$ be Borel probability measures in $\mathbf{R}$. The space $\mathbf{R}^{\mathbf{N}_+}$ is, by definition, the set of all sequences $x = (x_k)_{k=1}^{\infty}$

of real numbers. For each $k \in \mathbf{N}_+$, set $\pi_k(x) = x_k$. The σ-algebra $\mathcal{R}^{\mathbf{N}_+}$ is the least σ-algebra $\mathcal{S}$ of subsets of $\mathbf{R}^{\mathbf{N}_+}$ which makes all the projections $\pi_k : (\mathbf{R}^{\mathbf{N}_+}, \mathcal{S}) \to (\mathbf{R}, \mathcal{R})$, $k \in \mathbf{N}_+$, measurable. Below, $(\pi_1, ..., \pi_n)$ denotes the mapping of $\mathbf{R}^{\mathbf{N}_+}$ into $\mathbf{R}^n$ defined by the equation

$$(\pi_1, ..., \pi_n)(x) = (\pi_1(x), ..., \pi_n(x)).$$

Theorem 3.6.1. *There is a unique probability measure μ on $\mathcal{R}^{\mathbf{N}_+}$ such that*

$$\mu_{(\pi_1,...,\pi_n)} = \mu_1 \times ... \times \mu_n$$

for every $n \in \mathbf{N}_+$.

The measure μ in Theorem 3.6.1 is called the product of the measures μ_k, $k \in \mathbf{N}_+$, and is often denoted by

$$\times_{k=1}^{\infty} \mu_k.$$

PROOF OF THEOREM 3.6.1. Let $(\Omega, P, \mathcal{F}) = ([0, 1[\,, v_{1|[0,1[}, \mathcal{B}([0, 1[)$ and set

$$\eta(\omega) = \Sigma_{k=1}^{\infty} \frac{\xi_k^{(2)}(\omega)}{2^k},\ \omega \in \Omega.$$

We already know that $P_\eta = P$. Now suppose $(k_i)_{i=1}^{\infty}$ is a strictly increasing sequence of positive integers and introduce

$$\eta' = \Sigma_{i=1}^{\infty} \frac{\xi_{k_i}^{(2)}(\omega)}{2^i},\ \omega \in \Omega.$$

Note that for each fixed positive integer n, the $\mathbf{R}^n$-valued maps $(\xi_1^{(2)}, ..., \xi_n^{(2)})$ and $(\xi_{k_1}^{(2)}, ..., \xi_{k_n}^{(2)})$ are P-equimeasurable. Thus, if $f : \Omega \to \mathbf{R}$ is continuous,

$$\int_\Omega f(\eta) dP = \lim_{n\to\infty} \int_\Omega f(\Sigma_{k=1}^{n} \frac{\xi_k^{(2)}(\omega)}{2^k}) dP(\omega)$$

$$= \lim_{n\to\infty} \int_\Omega f(\Sigma_{i=1}^{n} \frac{\xi_{k_i}^{(2)}(\omega)}{2^i}) dP(\omega) = \int_\Omega f(\eta') dP$$

and it follows that $P_{\eta'} = P_\eta = P$.

By induction, we define for each $k \in \mathbf{N}_+$ an infinite subset N_k of the set $\mathbf{N}_+ \setminus \cup_{i=1}^{k-1} N_i$ such that the set $\mathbf{N}_+ \setminus \cup_{i=1}^{k} N_i$ contains infinitely many elements and define

$$\eta_k = \Sigma_{i=1}^{\infty} \frac{\xi_{n_{ik}}^{(2)}(\omega)}{2^i}$$

where $(n_{ik})_{i=1}^{\infty}$ is an enumeration of N_k. The map

$$\Psi(\omega) = (\eta_k(\omega))_{k=1}^{\infty}$$

is a measurable map of $(\Omega, \mathcal{F})$ into $(\mathbf{R}^{\mathbf{N}_+}, \mathcal{R}^{\mathbf{N}_+})$ and

$$P_\Psi = \times_{k=1}^{\infty} \lambda_i$$

where $\lambda_i = P$ for each $i \in \mathbf{N}_+$.

For each $i \in \mathbf{N}_+$ there exists a measurable map φ_i of $(\Omega, \mathcal{F})$ into $(\mathbf{R}, \mathcal{R})$ such that $P_{\varphi_i} = \mu_i$ (see Section 1.6). The map

$$\Gamma(x) = (\varphi_i(x_i))_{i=1}^{\infty}$$

is a measurable map of $(\mathbf{R}^{\mathbf{N}_+}, \mathcal{R}^{\mathbf{N}_+})$ into itself and we get $\mu = (P_\Psi)_\Gamma$. This completes the proof of Theorem 3.6.1.

$\uparrow\uparrow\uparrow$