Notes on Product Measures for Math 501

Eric A. Carlen[1]
Rutgers University

November 19, 2014

1 Product Algebras

Recall a lemma we proved earlier, as Lemma 6.1 in the notes on σ-algebras and measurability.:

1.1 LEMMA. *Let X be any set, and let $\mathcal{F}$ be a set of subsets of X such that for all E and F in $\mathcal{F}$, $E\backslash F$ is a finite disjoint union of elements of $\mathcal{F}$. Then the set $\mathcal{A}$ of all finite disjoint unions of elements of $\mathcal{F}$ is an algebra.*

1.2 DEFINITION. Let X be any set. An *elementary family* of subsets of X is any set $\mathcal{F}$ of subsets of X such that for all $E, F \in \mathcal{F}$, $E\backslash F$ is a finite disjoint union of sets in $\mathcal{F}$. For any set $\mathcal{F}$ of subsets of X, $\alpha(\mathcal{F})$ is the smallest algebra containing $\mathcal{F}$.

Lemma 1.1 says that if $\mathcal{F}$ is an elementary family, then $\alpha(\mathcal{F})$ consists of all finite disjoint unions of sets in $\mathcal{F}$.

1.3 DEFINITION. Now let X and Y be two sets. Let $\mathcal{F}$ and $\mathcal{G}$ be elementary families of sets in X and Y respectively. Then rectangle$(\mathcal{F},\mathcal{G})$ consists of all sets $E \subset X \times Y$ of the form $F \times G$ with $F \in \mathcal{F}$ and $G \in \mathcal{G}$.

1.4 LEMMA. *Let X and Y be two sets. Let $\mathcal{F}$ and $\mathcal{G}$ be elementary families of sets in X and Y respectively. Then* rectangle$(\mathcal{F},\mathcal{G})$ *is an elementary family of sets in $X \times Y$.*

Proof. Let $E_1 = F_1 \times G_1$ and $E_2 = F_2 \times G_2$ belong to $\mathcal{P}$. Then

$$E_2^c = (F_2^c \times G_2)\bigcup(F_2^c \times G_2^c)\bigcup(F_2 \times G_2^c) \ ,$$

and this union is disjoint. Therefore,

$$E_1\backslash E_2 = E_1 \cap E_2^c = (F_1\backslash F_2 \times G_1 \cap G_2)\bigcup(F_1\backslash F_2 \times G_1\backslash G_2)\bigcup(F_1 \cap F_2 \times G_1\backslash G_2) \ ,$$

By Lemma, 1.1, $F_1\backslash F_2$ and $G_1 \cap G_2$ are finite disjoint union of sets in $\mathcal{F}$ and $\mathcal{G}$ respectively. Hence $F_1\backslash F_2 \times G_1 \cap G_2$ is a finite disjoint union of sets in rectangle$(\mathcal{F},\mathcal{G})$. The same applies to $F_1\backslash F_2 \times G_1\backslash G_2$ and $F_1 \cap F_2 \times G_1\backslash G_2$. □

Of course, an algebra is an elementary family. Thus, as a special case, the lemma implies that if $\mathcal{A}$ is an algebra of subsets of X, and $\mathcal{B}$ is an algebra of subsets of Y, then the set of all finite disjoint unions of sets in rectangle$(\mathcal{A}, \mathcal{B})$ is an algebra of subsets of $X \times Y$. This justifies the following definition:

1.5 DEFINITION (Product algebra)**.** Let X and Y be two sets. Let $\mathcal{A}$ be an algebra of subsets of X, and $\mathcal{B}$ be an algebra of subsets of Y. Then the product algebra $\mathcal{A}_1 \otimes \mathcal{A}_2$ is the algebra consisting of all finite unions of sets in rectangle$(\mathcal{A}, \mathcal{B})$.

Combining results we have proved:

1.6 THEOREM. *Let X and Y be two sets. Let $\mathcal{F}$ and $\mathcal{G}$ be elementary families of sets in X and Y respectively, and let $\mathcal{A} = \alpha(\mathcal{F})$ and $\mathcal{B} = \alpha(\mathcal{G})$. Then $\mathcal{A} \otimes \mathcal{B}$ consists of all finite disjoint unions of sets of the form $F \times G$ with $F \in \mathcal{F}$ and $G \in \mathcal{G}$.*

2 Product σ-Algebras

2.1 DEFINITION. Let $\mathcal{M}$ and $\mathcal{N}$ be σ-algebras of subsets of X and Y respectively. A *rectangle* is a subset of $X \times Y$ of the form $F \times G$ with $F \in \mathcal{M}$ and $G \in \mathcal{N}$. Let $\mathcal{A}$ be the algebra consisting of all finite disjoint unions of rectangles. The *product σ-algebra* $\mathcal{M} \otimes \mathcal{N}$ is the smallest σ-algebra containing the algebra $\mathcal{A}$.

2.2 Remark. Our notation is somewhat ambiguous: Since σ-algebras are also algebras, $\mathcal{M} \otimes \mathcal{N}$ can be interpreted two ways. We shall always mean the product σ-algebra instead of the smaller product algebra unless we explicitly indicate otherwise.

2.3 DEFINITION. Let $E \subset X \times Y$. For all $y \in Y$, define

$$S_1(E, y) = \{x \in X \ : (x, y) \in E\} \ ,$$

and for all $x \in X$, define

$$S_2(E, x) = \{y \in Y \ : (x, y) \in E\} \ ,$$

These are the *slices of E through y and x respectively.*

It is immediate from the definition that if $\{E_n\}_{n \in \mathbb{N}}$ is any sequence of sets in $X \times Y$,

$$S_1\left(\bigcup_{n=1}^{\infty} E_n, y\right) = \bigcup_{n=1}^{\infty} S_1(E_n, y) \quad \text{and} \qquad S_1\left(\bigcap_{n=1}^{\infty} E_n, y\right) = \bigcap_{n=1}^{\infty} S_1(E_n, y) \tag{2.1}$$

for all $y \in Y$. Of course, the analogous formulas hold for S_2. Likewise, for any $E \subset X \times Y$, $x \in S_1(E^c, y)$ if and only if $(x, y) \in E^c$, which is true if and only if $(x, y) \notin E$, so that

$$S_1(E^c, y) = (S_1(E, y))^c \ , \tag{2.2}$$

and the analogous identity is valid for S_2. In short, taking of slices commutes with taking of complements.

2.4 THEOREM. *Let X and Y be sets, and let $\mathcal{A}$ and $\mathcal{B}$ be algebras of sets in X and Y. Then for all $E \in \sigma(\mathcal{A} \otimes \mathcal{B})$, and all $x \in X$, and all $y \in Y$, $S_1(E, y) \in \sigma(A_1)$ and $S_2(E, x) \in \sigma(A_2)$.*

Proof. Let $\mathcal{S}$ be the set of subsets of $X \times Y$ such that for all $E \in \mathcal{M}$, and all $x \in X$, and all $y \in Y$, $S_1(E, y) \in \sigma(A_1)$ and $S_2(E, x) \in \sigma(A_2)$. By (2.1) and (2.2), $\mathcal{S}$ is a σ-algebra. It remains to show that $\mathcal{A} \otimes \mathcal{B} \in \mathcal{S}$. In fact, since the general element of $\mathcal{A} \otimes \mathcal{B}$ is a disjoint union of rectangles, it suffices to show that these rectangles are contained in $\mathcal{S}$.

Let $E = F \times G$ with $F \in \mathcal{A}$ and $G \in \mathcal{B}$. Then

$$S_1(E, y) = \begin{cases} F & y \in G \\ \emptyset & y \notin G \end{cases} .$$

Thus, for all y, $S_1(E, y) \in \sigma(\mathcal{A})$, and the same reasoning applies to $S_2(E, x)$. Thus, all sets $E = F \times G$ with $F \in \mathcal{A}$ and $G \in \mathcal{G}$ belong to $\mathcal{S}$. □

3 Product measures

3.1 LEMMA. *Let X and Y be sets, and let $\mathcal{A}$ and $\mathcal{B}$ be algebras of sets in X and Y. Let m be a premeasure on $\mathcal{A}$, and let n be a premeasure on $\mathcal{B}$. Let $A \in \mathcal{A} \otimes \mathcal{B}$ have two representations as a disjoint union of rectangles:*

$$A = A = \bigcup_{j=1}^{N} F_j \times G_j \qquad \text{and} \qquad A = \bigcup_{k=1}^{M} \widetilde{F}_k \times \widetilde{G}_k \ .$$

Then

$$\sum_{j=1}^{M} m(F_j)n(G_j) = \sum_{k=1}^{M} m(\widetilde{F}_k)n(\widetilde{G}_k) \ .$$

Proof. Taking intersections we also have

$$A = \bigcup_{1 \le j \le N, 1 \le k \le M} (F_k \cap \widetilde{F_j}) \times (G_k \cap \widetilde{G_j}) \ .$$

By symmetry, it suffices to show that $\displaystyle\sum_{j=1}^{M} m(F_j)n(G_j) = \sum_{j=1}^{N} \sum_{k=1}^{M} m(F_k \cap \widetilde{F_j})n(G_k \cap \widetilde{G_j})$. For this, it suffices to show that for each j, $m(F_j)n(G_j) = \displaystyle\sum_{k=1}^{M} m(F_k \cap \widetilde{F_j})n(G_k \cap \widetilde{G_j})$

Simplifying our notation, it suffices to show that whenever $F \in \mathcal{A}$ and $G \in \mathcal{B}$, and

$$F \times G = \cup_{j=1}^{N} F_j \times G_j \ , \tag{3.1}$$

where the right hand side is a disjoint union of rectangles, then

$$m(F)n(G) = \sum_{j=1}^{N} m(F_j)n(G_j) \ .$$

To prove this, let S be any non-empty subset of $\{1,\dots,N\}$, and define $F_S = \left(\bigcap_{j\in S} F_j\right)\bigcap\left(\bigcap_{k\in S^c} F_j^c\right)$. By construction, for $S \neq S'$, $F_S \cap F_{S'} = \emptyset$. (It is also possible that $F_S = \emptyset$ for some S.) Also, for each j, $F_j = \bigcup_{\{S\ :\ j\in S\}} F_S$ and $F = \bigcup_S F_S$. We define G_S using the analogous construction. Then

$$F_j \times G_j = \bigcup_{\{S,S'\ :\ j\in S, j\in S'\}} F_S \times G_{S'}\ .$$

Since (3.1) is a *disjoint* union, for $j \neq k$,

$$\left(\bigcup_{\{S,S'\ :\ j\in S, j\in S'\}} F_S \times G_{S'}\right)\bigcap\left(\bigcup_{\{S,S'\ :\ k\in S, k\in S'\}} F_S \times G_{S'}\right) = \emptyset\ ,$$

and so whenever $S\bigcap S' \neq \{j\}$ for some $j \in \{1,\dots,N\}$, $F_S \times G_{S'} = \emptyset$. Thus,

$$S\bigcap S' \neq \{j\} \quad \text{for some} \quad j \in \{1,\dots,N\} \quad \Rightarrow m(F_s)n(G_{S'}) = 0 \tag{3.2}$$

where we have used the fact that m and n are premeasures so that $m(\emptyset) = n(\emptyset) = 0$. Since m and n are finitely additive,

$$\begin{aligned}
m(F)n(G) &= \left(\sum_{S\subset\{1,\dots,N\}} m(F_S)\right)\left(\sum_{S'\subset\{1,\dots,N\}} n(G_{S'})\right)\\
&= \sum_{S,S'} m(F_S)n(G_{S'})\\
&= \sum_{j=1}^{N}\left(\sum_{S,S'\ :S\cap S'=\{j\}} m(F_S)n(G_{S'})\right)\\
&= \sum_{j=1}^{N}\left(\sum_{S,S'\ :j\in S, j\in S'} m(F_S)n(G_{S'})\right)\\
&= \sum_{j=1}^{N}\left(\sum_{S\ \ j\in S} m(F_S)\right)\left(\sum_{S'\ \ j\in S'} n(G_{S'})\right)\\
&= \sum_{j=1}^{N} m(F_j)n(G_j)\ .
\end{aligned}$$

□

On the basis of Lemma 3.1, we may define a function $m\otimes n$ on $\mathcal{A}\otimes\mathcal{B}$ by

$$m\otimes n\left(\bigcup_{j=1}^{N} F_k \times G_k\right) = \sum_{j=1}^{M} m(F_j)n(G_j) \tag{3.3}$$

whenever $\{F_j \times G_j\}_{j=1,\dots,N}$ is a set of disjoint rectangles in $\mathcal{A}\otimes\mathcal{B}$. It is clear that $m\otimes n(\emptyset) = 0$, and that $m\otimes n$ is finitely additive on $\mathcal{A}\otimes\mathcal{B}$. Hence $m\otimes n$ is a premeasure on $\mathcal{A}\otimes\mathcal{B}$. It is called the *product of the premeasures m and n.*

3.2 LEMMA. *Let X and Y be sets, and let $\mathcal{A}$ and $\mathcal{B}$ be algebras of sets in X and Y. Let m be a premeasure on $\mathcal{A}$, and let n be a premeasure on $\mathcal{B}$. Suppose that both m and n are continuous at the empty set and are semifinite. Then the product premeasure $m \otimes n$ is continuous at the empty set and is semifinite.*

Proof. The proof that $m \otimes n$ is semifinite is left to the reader. Let $\{E_k\}_{k\in\mathbb{N}}$ be a sequence of sets in $\mathcal{A} \otimes \mathcal{B}$ such that $E_{k+1} \subset E_k$ for all k, $m \otimes n(E_1) < \infty$, and $\bigcap_{k=1}^{\infty} E_k = \emptyset$. We must show that

$$\lim_{k\to\infty} m \otimes n(E_k) = 0 \ .$$

To do this, fix $k \in \mathbb{N}$, and write E_k as a disjoint union of rectangles

$$E_k = \bigcup_{j=1}^{N} F_j \times G_j \ .$$

Then $m \otimes n(E_k) = \sum_{j=1}^{N} m(F_j)n(G_j)$, and $S_2(E_k, x) = \displaystyle\bigcup_{\{j\ :\ x\in F_j\}} G_j$ is a disjoint union. Consequently,

$$n(S_2(E_k, x)) = \sum_{j=1}^{N} n(G_j)1_{E_j}(x) \ .$$

Let μ be the countably additive extension of m to $\sigma(\mathcal{A})$, and let g_k be the simple function given by $g_k(x) = n(S_2(E_k, x))$. Then, since μ extends m,

$$\int_X g_k \mathrm{d}\mu = \sum_{j=1}^{N} n(G_j) \int_X 1_{E_j} \mathrm{d}\mu = \sum_{j=1}^{N} n(G_j)m(E_j) = m \otimes n(E_k) \ . \tag{3.4}$$

Since k is arbitrary, we have $\displaystyle\lim_{k\to\infty} m \otimes n(E_k) = \lim_{k\to\infty} \int_X g_k \mathrm{d}\mu$. Since $\bigcap_{k=1}^{\infty} E_k = \emptyset$, for each x, $\bigcap_{k=1}^{\infty} S_2(E_k, x) = \emptyset$, and then since n is continuous at the empty set,

$$\lim_{k\to\ infty} g_k(x) = \lim_{k\to\ infty} n\left(S_2(E_k, x)\right) = 0 \ .$$

Thus, $\lim_{k\to\infty} g_k(x) = 0$ for all x. Furthermore $g_k \le g_1$, and $\int_X g_1 \mathrm{d}\mu = m \otimes n(E_1) < \infty$. Hence, by the Lebesgue Dominated Convergence Theorem,

$$\lim_{k\to\infty} \int_X g_k \mathrm{d}\mu = 0 \ .$$

By (3.4), this completes the proof. □

3.3 DEFINITION (The product measure $\mu \otimes \nu$)**.** Let X and Y be sets, and let $\mathcal{M}$ and $\mathcal{N}$ be σ algebras of sets in X and Y.Let μ and ν be σ-finite measures on $\mathcal{M}$ and $\mathcal{N}$ respectively. Let $\mathcal{M} \otimes \mathcal{N}$ be the algebraic product of $\mathcal{M}$ and $\mathcal{N}$ regarded as algebras. Regard μ and ν as σ-finite premeasures on $\mathcal{M}$ and $\mathcal{N}$ respectively, and let m be the corresponding product premeasure. Since by Lemma 3.2, m is semifinite and continuous at the empty set, m has a unique countably additive extension to the σ-algebra $\mathcal{M} \otimes \mathcal{N}$, and by the σ-finiteness of m, this extension is unique. This measure is the *product measure* $\mu \otimes \nu$ on $\mathcal{M} \otimes \mathcal{N}$.

By the definition, for all $F \in \mathcal{M}$ and all $G \in \mathcal{N}$, $\mu \otimes \nu(F \times G) = \mu(F)\nu(G)$, and, by the uniqueness theorem, which applies on account of our assumption that μ and ν are σ-finite, $\mu \otimes \nu$ is the unique measure on $\mathcal{M} \otimes \mathcal{N}$ with this property. If we only assume that μ and ν are semifinite, then we can still define $\mu \otimes \nu$ on $\mathcal{M} \otimes \mathcal{N}$ using the Caratheodory construction, but then uniqueness may fail.

4 The Fubini-Tonelli Theorem

Let f be a function on $X \times Y$. For $y \in Y$, $f(\cdot, y)$ denotes the function on X whose value at x is $f(x, y)$. Likewise, for $x \in X$, $f(x, \cdot)$ denotes the function on Y whose value at x is $f(x, y)$.

4.1 LEMMA. *Let f be a measurable function on $(X \times Y, \mathcal{M} \otimes \mathcal{N})$. Then for all $y \in Y$, $f(\cdot, y)$ is a measurable function on $(X, \mathcal{M})$, and for all $x \in X$, $f(x, \cdot y)$ is a measurable function on $(Y, \mathcal{N})$.*

Proof. Suppose f is a real valued measurable function on $(X \times Y, \mathcal{M} \otimes \mathcal{N})$ Then for all $a \in \mathbb{R}$, $f^{-1}((a, \infty)) \in \mathcal{M} \otimes \mathcal{N}$, Then since

$$f^{-1}(\cdot, y)((a, \infty)) = S_1\left(f^{-1}((a, \infty))\ ,\ y\right)\ ,$$

and since the right side belongs to $\mathcal{M}$ by what we have said above and Theorem 2.4, $f^{-1}(\cdot, y)((a, \infty)) \in \mathcal{M}$. Since this is true for all $a \in \mathbb{R}$, $f(\cdot, y)$ is measurable on $(X, \mathcal{M})$. The same reasoning applies to $f(x, \cdot)$. □

Lemma 4.1 tells us that when $f \in L^+(X \times Y, \mu \otimes \nu)$, then the functions g and h given by

$$g(x) = \int_Y f(x, \cdot)\mathrm{d}\nu \qquad \text{and} \qquad h(y) = \int_X f(\cdot, y)\mathrm{d}\mu \tag{4.1}$$

are well-defined on X and Y respectively. In fact, much more is true:

4.2 THEOREM (Tonelli's Theorem)**.** *Let X and Y be sets, and let $\mathcal{M}$ and $\mathcal{N}$ be σ algebras of sets in X and Y.Let μ and ν be σ-finite measures on $\mathcal{M}$ and $\mathcal{N}$ respectively. Let $\mu \otimes \nu$ be their product measure on $\mathcal{M} \otimes \mathcal{N}$. Let $f \in L^+(X \times Y, \mu \otimes \nu)$ and let g and h be given by (4.1). Then $g \in L^+(X, \mathcal{M})$ and $h \in L^+(Y, \mathcal{N})$, and*

$$\int_X g\mathrm{d}\mu = \int_{X \times Y} f\mathrm{d}\mu \otimes \nu = \int_Y h\mathrm{d}\nu\ . \tag{4.2}$$

Proof. Let $E = F \times G$ with $F \in \mathcal{M}$ and $G \in \mathcal{N}$, and consider the case in which $f = 1_E$. Then with g and h defined by (4.1), $g(x) = \nu(G)1_F(x) \in L^+(X, \mathcal{M})$ and $h(y) = \mu(F)1_G(y) \in L^+(Y, \mathcal{N})$. It is now evident that

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

so that (4.2) is valid in this case. From the additivity and homogeneity of integration it follows that whenever $f = 1_E$ and E is a finite disjoint union of rectangles, $g(x) \in L^+(X, \mathcal{M})$, $h(y) \in L^+(Y, \mathcal{N})$ and (4.2) is valid. Recall that the set of such sets is an algebra, namely the algebraic product of $\mathcal{M}$ and $\mathcal{N}$.

Let $\mathcal{S}$ be the sets of all sets E in the σ-algebra $\mathcal{M}\otimes\mathcal{N}$ such that when $f = 1_E$, and g and h are given in terms of f by (4.1), then $g(x) \in L^+(X,\mathcal{M})$, $h(y) \in L^+(Y,\mathcal{N})$ and (4.2) is valid. We claim that $\mathcal{S}$ is a monotone class. One this is shown, it will follow from the Monotone Class Theorem and the fact that the algebra consisting of all finite disjoint unions of rectangles is contained in $\mathcal{S}$ that $\mathcal{S}$ is the σ-algebra generated by this algebra. But that is precisely $\mathcal{M}\otimes\mathcal{N}$, and so we will have that $\mathcal{S} = \mathcal{M}\otimes\mathcal{N}$.

Let $\{E_n\}_{n\in\mathbb{N}}$ be an increasing sequence of sets in $\mathcal{S}$. Let $f_n = 1_{E_n}$ and let g_n and h_n be given in terms of f_n by (4.1). Let $E = \bigcup_{n=1}^\infty E_n$. Then by continuity from below,

$$\mu\otimes\nu(E) = \lim_{n\to\infty}\mu\otimes\nu(E_n) = \lim_{n\to\infty}\int_X g_n(x)\mathrm{d}\mu\ . \tag{4.3}$$

Then since $g_n(x) = \nu(S_2(E_n,x))$, and since $\{S_2(E_n,x)\}_{n\in\mathbb{N}}$ is a sequence of sets increasing to $S_2(E,x)$, another application of continuity from below yields the fact that $\{g_n\}_{n\in\mathbb{N}}$ increases to $g(x)$ for all x. Then by the Lebesgue Monotone Convergence Theorem,

$$\lim_{n\to\infty}\int_X g_n\mathrm{d}\mu = \int_X g\mathrm{d}\mu\ .$$

Combining this with (4.3), we have that $g \in L^+(X,\mathcal{M})$ and

$$\mu\otimes\nu(E) = \int_{X\times Y} f\mathrm{d}\mu\otimes\nu = \int_X g\mathrm{d}\mu\ .$$

The same sort of reasoning shows that $h \in L^+(Y,\mathcal{N})$ and $\mu\otimes\nu(E) = \int_Y h\mathrm{d}\nu$.

Now we must show that if $\{E_n\}_{n\in\mathbb{N}}$ is a decreasing sequence of sets in $\mathcal{S}$, and $E = \bigcap_{n=1}^\infty E_n$, then $E \in \mathcal{S}$. The argument is essentially the same as the one we have just given or increasing sequences except that continuity from above will replace continuity from below, and the Lebesgue Dominated Convergence Theorem will replace the Lebesgue Monotone Convergence Theorem. Both of these require finiteness conditions. When $\mu(X)$ and $\nu(Y)$ are both finite, so that $\mu\otimes\nu(X\times Y) < \infty$, these finite conditions are trivially satisfied. The σ-finite case reduces to the finite case in the standard way, and the details are left to the reader.

At this point we have shown that for all $E \in \mathcal{M}\otimes\mathcal{N}$, with $f = 1_E$ and g and h given in terms of f by (4.1), then $g(x) \in L^+(X,\mathcal{M})$, $h(y) \in L^+(Y,\mathcal{N})$ and (4.2) is valid. By the additivity and homogeneity of integration, the same is true for all simple functions $f \in L^+(X\times Y,\mathcal{M}\otimes\mathcal{N})$.

Finally, we use the fact that every $f \in L^+(X\times Y,\mu\otimes\nu)$ is the pointwise limit of an increasing sequence $\{f_n\}_{n\in\mathbb{N}}$ of simple functions. But then for each $y \in Y$, $\{f_n(\cdot,y)\}_{n\in\mathbb{N}}$ is an sequence of simple functions converging pointwise to $f(\cdot,y)$. By Lebesgue's Monotone Convergence Theorem,

$$h(y) = \int_X f(\cdot,y)\mathrm{d}\mu = \lim_{n\to\infty}\int_X f_n(\cdot,y)\mathrm{d}\mu\ .$$

Let $h_n(y) = \int_X f_n(\cdot,y)\mathrm{d}\mu$. By what we have noted above, $h_n \in L^+(Y,\mathcal{N})$. Since $L^+(Y,\mathcal{N})$ is closed under pointwise limits, $h \in L^+(Y,\mathcal{N})$. Similar reasoning shows that $g \in L^+(X,\mathcal{M})$. □

4.3 THEOREM (The Fubini-Tonelli Theorem). *Let X and Y be sets, and let $\mathcal{M}$ and $\mathcal{N}$ be σ algebras of sets in X and Y.Let μ and ν be σ-finite measures on $\mathcal{M}$ and $\mathcal{N}$ respectively. Let $\mu\otimes\nu$*

be their product measure on $\mathcal{M}\otimes\mathcal{N}$. Let f be measurable functions on $(X\times Y,\mu\otimes\nu)$ Then each of the following integral are well defined:

$$\int_{X\times Y}|f(x,y)|\mathrm{d}\mu\otimes\nu \tag{4.4}$$

$$\int_X\left(\int_Y|f(x,y)|\mathrm{d}\nu\right)\mathrm{d}\mu \tag{4.5}$$

$$\int_Y\left(\int_X|f(x,y)|\mathrm{d}\mu\right)\mathrm{d}\nu \tag{4.6}$$

Moreover, if any one of them is finite, then all are finite and equal and

$$\int_{X\times Y}f(x,y)\mathrm{d}\mu\otimes\nu=\int_X\left(\int_Y f(x,y)\mathrm{d}\nu\right)\mathrm{d}\mu=\int_Y\left(\int_X|f(x,y)|\mathrm{d}\mu\right)\mathrm{d}\nu\ . \tag{4.7}$$

Proof. Since $|f|\in L^+(X\times Y,\mu\otimes\mathcal{N})$, the first part follows from Tonelli's Theorem. Then, assuming that any one of the three integrals (4.4), (4.5) or (4.6) is finite, then $f\in L^1(X\times Y,\mathcal{M}\otimes\mathcal{N},\mu\otimes\nu)$, and now decomposing f in to the positive and negative parts of its real and imaginary parts, and applying Tonelli's Theorems to each of these, we obtain the final part. □

4.4 Remark. In 1906 Fubini's proved the above theorem except that he required that the integral (4.4) was finite. In 1909 Tonelli showed how one could verify Fubini's condition – finiteness of (4.4) – by checking finiteness of an iterated integral. However, it is still essential that f be *measurable* with respect to the product σ-algebra $\mathcal{M}\otimes\mathcal{N}$. It is *not* sufficient that, say $f(x,\cdot)\in L^1(Y,\mathcal{N},\nu)$ for each x, and that $x\mapsto\int_Y|f(x,y)|\mathrm{d}\nu\in L^1(X,\mathcal{M},\mu)$. This would be enough to guarantee that the integral in (4.5) is well defined and finite. But unless f is measurable with respect to $\mathcal{M}\otimes\mathcal{N}$, the integral in (4.4) is not even defined and the integrals in (4.7) need not be equal.

5 Construction of Lebesgue measure on $\mathbb{R}^n$

For $n\in\mathbb{N}$, define $\mathcal{E}_n$ to be the set of half open rectangles in $\mathbb{R}^n$; i.e., the sets of the form

$$\{x\in\mathbb{R}^n\ :\ \langle\mathbf{e}_j,x\rangle\in(a_j,b_j]\quad j=1,\dots,n\ \} \tag{5.1}$$

where $\mathbf{e}_j$ is the jth standard basis vector in $\mathbb{R}^n$, and where $\langle x,y\rangle$ denotes the standard inner product in $\mathbb{R}^n$. Since $\mathcal{E}_n$ is an elementary family, $\mathcal{A}_n$, the set of all finite disjoint unions of sets in $\mathcal{E}_n$ is an algebra.

For $E\in\mathcal{E}_n$, given by (5.1) define

$$\rho_n(E)=\prod_{j=1}^n(b_j-a_j)\ . \tag{5.2}$$

Let $A\in\mathcal{A}_n$. By definition, A is the disjoint of finitely many half-open rectangles $E_1,\dots,E_N$. We then define

$$\rho_n(A)=\sum_{j=1}^N\rho_n(E_j)\ ,$$

and note that there is no ambiguity stemming from the fact that A can be written in more than one way as a finite disjoint union of half-open rectangles: Considering any common refinement of two such partitions of A into half-open rectangles, we see that they yield the same value for $\rho(A)$.

5.1 THEOREM. *For each $n \in \mathbb{N}$, ρ_n is a semi-finite premeasure on $\mathcal{A}_n$ that is continuous at the empty set.*

Proof. The fact the ρ_n is semifinite is clear. We prove that it is continuous at the empty set by induction on n, noting that we have already proved this for $n = 1$.

Suppose we have shown that ρ_m is continuous at the empty set for all $m < n$. One readily checks that $\mathcal{A}_n = \mathcal{A}_1 \otimes \mathcal{A}_{n-1}$ and that $\rho_n = \rho_1 \otimes \rho_{n-1}$. Then by Lemma 3.2 and our inductive hypothesis, ρ_n is continuous at the empty set. □

5.2 DEFINITION (Lebesgue outer measure and measure)**.** The outer measure μ_n^* on $\mathbb{R}^n$ defined by

$$\mu_n^*(E) = \inf \left\{ \sum_{j=1}^{\infty} \rho_n(A_j) \ : \ \{A_j\}_{j\in\mathbb{N}} \subset \mathcal{A}_n \quad \text{and} \quad E \subset \bigcup_{j=1}^{\infty} A_j \right\} \tag{5.3}$$

is the *Lebesgue outer measure on* $\mathbb{R}^n$. Its Caratheodory σ-algebra is the σ-algebra $\mathcal{L}_n$ of *Lebesgue measurable subsets of* $\mathbb{R}^n$, and the restriction m_n of μ_n^* restriction to $\mathcal{L}_n$ is Lebesgue measure on $\mathbb{R}^n$.

5.3 Remark. By decomposing each A_j into a finite union of half open rectangles, we see that we do not raise the infimum if we further require that each A_j in (5.3) is a half open rectangle. Next, slightly enlarging each of these, we see that we do not change the infimum if we further require that E is contained in the union of the interiors of the A_j. Extending the definition of ρ_n to open rectangles in the obvious way, we then have the alternate formula

$$\mu_n^*(E) = \inf \left\{ \sum_{j=1}^{\infty} \rho_n(A_j) \ : \ \text{each } A_j \text{ an open rectangle and} \quad E \subset \bigcup_{j=1}^{\infty} A_j \right\} \tag{5.4}$$

Because ρ_n is semifinite and continuous at the empty set, the restriction of μ_n^* to $\mathcal{A}_n$ agrees with ρ_n. We also know that $\sigma(\mathcal{A}_n) \subset \mathcal{L}_n$.

5.4 PROPOSITION. *$\sigma(\mathcal{A}_n) = \mathcal{B}_n$, the Borel σ-algebra on $\mathbb{R}^n$.*

Proof. Since every open rectangle in $\mathbb{R}^n$ is a countable union of half-open rectangles in $\mathbb{R}^n$, $\sigma(\mathcal{A}_n)$ contains all open rectangles in $\mathbb{R}^n$. Since every open set in $\mathbb{R}^n$ is a countable union of open rectangles, $\sigma(\mathcal{A}_n)$ contains all open sets, and hence contains $\mathcal{B}_n$. Conversely, since each half open rectangle is a Borel set, $\sigma(\mathcal{A}_n) \subset \mathcal{B}_n$. □

Now let $E \in \mathcal{L}_n$, and $\epsilon > 0$. Then there is a sequence of open rectangles $\{A_j\}$ such that with $U := \cup_{j=1}^{\infty} A_j$, so that U is open,

$$E \subset U \qquad \text{and} \quad m_n(U) \leq \sum_{j=1}^{\infty} m_n(A_j) = \sum_{j=1}^{\infty} \rho_n(A_j) \leq m_n(E) + \epsilon \ .$$

This shows that m_n is *outer regular* on $\mathcal{L}$.

We next show that m_n is *inner regular*: Let $C_N = \{\ x \ : \ \langle \mathbf{e}_j, x\rangle \in [-N, N] \ , \ j = 1, \ldots n \ \}$ be the centered closed cube of side length $2N$.

For any $E \in \mathcal{L}_n$, and any $\epsilon > 0$, let U be an open set containing $C_N \cap E^c$ such that $m_n(U) \leq m_n(C_N \cap E^c) + \epsilon$. Define $K_N := C_N \cap U^c$. Then K_N is compact and contained in $E \cap C_N$. Then

$$\begin{aligned} m_n(C_N) &= m_n(K_N) + m_n(C_N \cap U) \\ &\leq m_n(K_N) + m_n(C_N \cap E^c) + \epsilon \ . \end{aligned}$$

Therefore,

$$m_n(K_N) \geq m_n(C_N) - m_n(C_N \cap E^c) - \epsilon = m_n(C_N \cap E) - \epsilon \ .$$

Now suppose $m_n(E) < \infty$. Then by continuity from below, there exists N so that $m_n(C_N \cap E) \geq m_n(E) - \epsilon$. Then with K_N as above, $m_n(K_N) \geq m_n(E) - 2\epsilon$. On the other hand, if $m_n(E) = \infty$, then for each k, there is an N so that $m_n(C_N \cap E) \geq k + 1$, and then, taking $\epsilon = 1$, $m_n(K_N) \geq k$. Either way, we have that

$$m_n(E) = \sup\{ \ m_n(K) \ : \ K \text{ compact}, K \subset E \ \} \ .$$

This shows that m is *inner regular* on $\mathcal{L}$. We summarize and extend:

5.5 THEOREM. *Lebesgue measure m_n is inner and outer regular on the Lebesgue measurable subsets of $\mathbb{R}^n$. Moreover, for every E in $\mathcal{L}_n$, there are Borel sets F and G such that*

$$F \subset E \subset G \qquad \text{and} \qquad m_n(G \cap F^c) = 0 \ .$$

In fact, we can take F to be a a countable union of closed sets, and G to be a countable intersection of open sets.

Proof. Write $\mathbb{R}^n$ as the disjoint union of a family of bounded half open rectangles $\{C_j\}$, which we may as well take to be cubes of unit side length. Given $k \in \mathbb{N}$, for each j there exists an open set $U_{j,k}$ such that $E \cap C_j \subset U_{j,k}$ and

$$m_n(U_{j,k}) \leq m_n(E \cap C_j) + \frac{1}{k2^j} \ .$$

Let $U_k = \cup_{j=1}^{\infty} U_{j,k}$ which is open. Then $E \subset U_k$ and

$$\mu(U_k \cap E^c) \leq \sum_{j=1}^{\infty} m_n(U_{j,k} \cap E^c) \leq \sum_{j=1}^{\infty} m_n(U_{j,k} \cap (E \cap C_j)^c) \leq \frac{1}{k} \ .$$

Define $G = \cap_{k=1}^{\infty} U_k$. Then G is a G_δ set, hence Borel, $E \subset G$, and $m_n(G \cap E^c) = 0$.

Likewise, choose $F_{j,k}$ to be a compact set contained in $E \cap C_j$ so that

$$m_n(F_{j,k}) \geq m_n(E \cap C_j) - \frac{1}{k2^j} \ .$$

Let $F = \cup_{j,k=1}^{\infty} F_{j,k}$. Then $F \subset E$ and F is an F_σ set. Finally,

$$m_n(E \cap F^c) = \sum_{j=1}^{\infty} m_n((E \cap C_j) \cap F^c) \leq \sum_{j=1}^{\infty} m_n((E \cap C_j) \cap F_{j,k}^c) \leq \frac{1}{k} \ .$$

Since this is true for all k, $m_n(E \cap F^c) = 0$. □

5.6 THEOREM (Approximation of measurable sets by finite unions of rectangles)**.** *Let $E \in \mathcal{L}_n$ be such that $m_n(E) < \infty$. Then for all $\epsilon > 0$, there is a set $A \in \mathcal{A}_n$ such that*

$$m_n(A\Delta E) \le \epsilon \ . \tag{5.5}$$

Proof. We have already seen, using the Monotone Class Theorem, that for any σ-finite measure μ on the σ algebra $\sigma(\mathcal{A})$ generated by some algebra $\mathcal{A}$, the following is true: For any measurable set E with $\mu(E) < \infty$, and any $\epsilon > 0$, there is an $A \in \mathcal{A}$ so that $\mu(A\Delta E) \le \epsilon$. Since the Borel σ-algebra of $\mathbb{R}^n$ is generated by $\mathcal{A}_n$ and since m_n is σ-finite, it follows that (5.5) is true whenever E is a Borel set of finite Lebesgue measure. But since we have shown above that every Lebesgue measurable set differs form a Borel set by a set of measure zero, the theorem is true in general. □

5.7 THEOREM. *$L^1(\mathbb{R}^n, \mathcal{L}_n, m_n)$ is separable and $C_c^\infty(\mathbb{R}^n)$ is dense in it.*

Proof. Let $f \in L^1(\mathbb{R}^n, \mathcal{L}_n, m_n)$, and let $\epsilon > 0$ be given. We know from the general theory of Lebesgue integration that there is a simple function ϕ such that $\|f - \phi\|_1 \le \epsilon$. Let $\phi = \sum_{j=1}^M z_j 1_{E_j}$. Without loss of generality, we may assume each z_j is rational. Approximating each E_j by a disjoint union of half-open rectangles, that we may assume to have rational boundaries, we find a function ψ of the form $\psi = \sum_{k=1}^N z_k 1_{R_k}$ where each z_k is rational, and each R_k is a half-open rectangle with rational boundary, and $\|\phi - \psi\|_1 \le \epsilon$. By Minkowski's inequality, $\|f - \psi\|_1 \le \|f - \phi\|_1 + \|\phi - \psi\|_2 \le 2\epsilon$. since there are only countably many functions of the form of ψ, $L^1(\mathbb{R}^n, \mathcal{L}_n, m_n)$, is separable.

To see that $C_c^\infty(\mathbb{R}^n)$ is dense, consider the function

$$\varphi(t) = \begin{cases} \exp(-(1-t^2)^{-1}) & -1 < t < 1 \\ 0 & |t| \ge 1 \ . \end{cases}$$

Note that $\varphi^{1/n}(t)$ increases monotonically with n to $1_{(-1,1)}(t)$. It follows that for any $a < b$ in $\mathbb{R}$, $\varphi^{1/k}((2t - (a+b))/(b-a))$ increases monotonically to $1_{(a,b)}(t)$ as $k \to \infty$, and hence that

$$\prod_{j=1}^n \varphi^{1/k}((2\langle \mathbf{e}_j, x\rangle - (a_j + b_j))/(b_j - a_j)) \quad \uparrow \quad \prod_{j=1}^n 1_{(a_j,b_j)}(\langle \mathbf{e}_j, x\rangle) \ .$$

Hence, by the Lebesgue Monotone Convergence Theorem, the characteristic function of any open rectangle, and hence any half-open rectangle, can be arbitrarily closely approximated in $L^1(\mathbb{R}^n, \mathcal{L}_n, m_n)$ by $C_c^\infty(\mathbb{R}^n)$ functions. Combining this with our approximation of f by a finite linear combination of characteristic functions of open rectangles, we obtain the result. □

6 Transformation of measures

Let $(X, \mathcal{M}, \mu)$ be a measure space, and let $(Y, \mathcal{N})$ be a measurable space. Let $T : X \to Y$ be measurable. Define a function $T\#\mu$ on $\mathcal{N}$ with values in $[0, \infty]$ by

$$T\#\mu(F) = \mu(T^{-1}(F)) \ .$$

Notice that if $\{F_j\}_{j\in\mathbb{N}}$ is disjoint in Y, then $\{T^{-1}(F_j)\}_{j\in\mathbb{N}}$ is disjoint in X and $T^{-1}(\cup_{j=1}^\infty F_j) = \cup_{j=1}^\infty T-1(F_j)$,

$$\begin{aligned} T\#(\cup_{j=1}^\infty(F_j)) &= \mu(T^{-1}(\cup_{j=1}^\infty F_j) = \mu(\cup_{j=1}^\infty T-1(F_j) \\ &= \sum_{j=1}^\infty \mu(T^{-1}(F_j)) = \sum_{j=1}^\infty T\#\mu(F_j)\ . \end{aligned}$$

Thus, $T\#\mu$ is a countably additive measure on $\mathcal{N}$. it is called the *push-forward* of μ by T. The following identity is the root of a number of *change of variables* formulae.

6.1 THEOREM. *Let $f \in L^+(Y,\mathcal{N})$, and let $T : X \to Y$ be measurable with respect to $\mathcal{N}$ on Y and $\mathcal{M}$ on X. Let μ be any measure on X, and define*

$$\nu = T\#\mu\ .$$

Then $f \circ T \in L^+(X,\mathcal{M})$, and

$$\int_X f\circ T\mathrm{d}\mu = \int_Y f\mathrm{d}\nu\ .$$

Proof. Clearly $f \circ T$ is non-negative and is measurable, so $f \circ T \in L^+(X,\mathcal{M})$, and both integrals are defined. Consider the case $f = 1_F, \in \mathcal{N}$. Then

$$\int_Y 1_F\mathrm{d}\nu = \nu(F) = \mu(T^{-1}(F_j)) = \int_X 1_{F_j}\circ T(x)\mathrm{d}\mu\ .$$

By linearity of integration, whenver f is a simple function, say $f = \sum_{j=1}^M z_j 1_{F_j}$ where each F_j belongs to $\mathcal{N}$.

$$\int_Y fT\mathrm{d}\nu = \int_X f\circ T\mathrm{d}\mu\ .$$

This proves the result for simple functions, and now the general result follows from approximation by simple functions. □

This theorem becomes useful if it can be combined with a concrete description of $T\#\mu$. We now turn to severla cases in whihc we can identify $T\#mu$ explicitly. In the first example, T is translation on $\mathbb{R}^n$, and we will show this leave Lebesgue measure invariant.

6.2 THEOREM (Translation invariance of Lebesgue measure)**.** *For all $a \in \mathbb{R}^n$, define $\tau_a : \mathbb{R}^n \to \mathbb{R}^n$ by $\tau(x) = x + a$. Then*

$$\tau_a\#m_n = m_n\ .$$

Proof. Let $A \in \mathcal{A}_n$. Note that τ_a is invertible and $\tau_a^{-1} = \tau_{-a}$. Notice that $\tau_a^{-1}(A) \in \mathcal{A}$. By definition,

$$\tau_a\#m_n(A) = m_n(\tau_a^{-1}(A)) = \rho_n(\tau_a^{-1}(A))\ ,$$

since on $\mathcal{A}_n$, m_n agrees with ρ_n. But ρ_n is invariant under translations as an obvious consequence of the formula that defines it. Hence

$$\rho_n(\tau_a^{-1}(A)) = \rho_n(A) = m_n(A)$$

for all $A \in \mathcal{A}_n$. Thus, $\tau_a\#m_n$ and m_n agree on $\mathcal{A}_n$, and so by the uniqueness theorem, they agree on all Borel sets. Finally, since every Lebesgue measurable contains and is contained in a Borel set of the same measure, the result is true for all $A \in \mathcal{L}_n$ as well. □

7 The push-forward of Lebesgue measure under an invertible linear transformation

7.1 THEOREM. *Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be linear and invertible. Then*

$$T\#m_n = |\det(T)|^{-1} m_n \ . \tag{7.1}$$

Consequently, for $f \in L^+(\mathbb{R}^n, \mathcal{L}_n, m_n)$,

$$|\det T| \int_{\mathbb{R}^n} f \circ T \mathrm{d}m_n = \int_{\mathbb{R}^n} f \mathrm{d}m_n \ . \tag{7.2}$$

Proof. The second part follows from the first part and Theorem 6.1. Hence it suffices to prove (7.1). The proof of this rests on the following elementary fact from linear algebra: Let E be any parallelepiped in $\mathbb{R}^n$. That is, E is some translate of the image of the unit cube in $\mathbb{R}^n$ under some invertible linear transformation $T : \mathbb{R}^n \to \mathbb{R}^n$. Then for every $\epsilon > 0$, there are two sets F and G in $\mathcal{A}_n$ such that $F \subset E \subset G$ and

$$\rho_n(G) - \epsilon \le |\det(T)| \le \rho_n(F) + \epsilon \ .$$

Since

$$\rho_n(F) \le m_n(E) \le m_n(G) \ ,$$

it follows that

$$m_n(E) = |\det(T)| \ .$$

Now let $T : \mathbb{R}^n \to \mathbb{R}^n$ be linear and invertible. Let S denote the inverse of T. Then S is also linear and has a matrix representation. Let $\mathbf{v}_j$ denote the jth row of this matrix so that

$$S(x) = (\langle \mathbf{v}_1, x\rangle, \dots, \langle \mathbf{v}_n, x\rangle) \ . \tag{7.3}$$

Since S is invertible, $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ is linearly independent, and hence a basis for $\mathbb{R}^n$. Conversely, any basis $\beta = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ generates an invertible linear transformation S on $\mathbb{R}^n$ through (7.3).

Now let

$$E := \{x \in \mathbb{R}^n \ : \ \langle \mathbf{e}_j, x\rangle \in (a_j, b_j] \quad j = 1, \dots, n \ \} \tag{7.4}$$

be a generic element of $\mathcal{E}_n$. Then

$$T(E) = S^{-1}(E) = \{x \in \mathbb{R}^n \ : \ \langle \mathbf{v}_j, x\rangle \in (a_j, b_j] \quad j = 1, \dots, n \ \} \ ,$$

is a rectangle in coordinates based on the basis $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$. It is evident that the sets of this form are an elementary family, and so the set of all finite disjoint unions of such sets forms an algebra that we denote $\mathcal{A}_T$. It is clear that T sets up a one-to-one correspondence between $\mathcal{A}_T$ and $\mathcal{A}_n$. That is, $E \in \mathcal{A}_n$ if and only if $T(A) \in \mathcal{A}_T$.

We define a premeasure ρ_T on $\mathcal{A}_T$ by

$$\rho_T(A) = \rho_n(T^{-1}(A)) \tag{7.5}$$

for all $A \in \mathcal{A}_T$, This is well defined since $T(A) \in \mathcal{A}_n$ for all $A \in \mathcal{A}_T$. Since ρ_n is σ-finite, and hence semifinite, so is ρ_T.

Moreover, if $\{A_j\}$ is a decreasing sequence with $\cap_{j=1}^{\infty} A_j = \emptyset$, then $\cap_{j=1}^{\infty} T^{-1}(A_j) = \emptyset$. It follows from the fact that ρ_n is continuous at the empty set that ρ_T has this property too.

It is clear that just as with $\mathcal{A}_n$, $\sigma(\mathcal{A}_T) = \mathcal{B}_n$. Let m_T denote the Borel measure on $\mathbb{R}^n$ obtained by restricting the outer measure generated by ρ_T to $\mathcal{B}_n$. Since, as noted at the beginning of the proof, the Lebesgue measure of every parallelepiped in $\mathbb{R}^n$ may be computed by evaluating an appropriate determinant, and since m_n extends ρ_n, it follows from (7.5) that

$$T\#m_n(A) = \rho_n(T^{-1}(A)) = |\det(T)|^{-1} m_n(A) \ ,$$

and hence that for all $A \in \mathcal{A}_T$,

$$T\#m_n(A) = m_T(A) = \rho_T(A) = |\det(T)|^{-1} m_n(A) \ , \tag{7.6}$$

Since m_T and $|\det(T)|m_n$ agree on the algebra $\mathcal{A}_T$, and since this algebra generates $\mathcal{B}_n$, it follows from the uniqueness theorem for measures agreeing on an algebra that $T\#m_n = m_T = |\det(T)|^{-1} m_n$ on all of $\mathcal{B}_n$. □

7.2 THEOREM (Rotation invariance of Lebesgue measure)**.** *Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be any orthogonal transformation. Then*

$$T\#m_n = m_n \ .$$

Proof. If T is orthogonal, $|\det(T)| = 1$, and the result follows from the previous theorem. □

7.3 THEOREM (Transformation of Lebesgue measure under dilation)**.** *For $t \in \mathbb{R}$ define $\varsigma_t : \mathbb{R}^n \to \mathbb{R}^n$ by $\varsigma_t(x) = e^{-t}x$. Then*

$$\varsigma_t\#m_n = e^{nt} m_n \ .$$

Proof. The transformation ς_t is linear and the corresponding matrix is e^{-t} times the identity, whose determinant is e^{-tn}. Hence this result is a direct consequence of (7.1). □

7.4 Remark. The set of transformations ς_t, $t \in \mathbb{R}$, form an abelian group of transformation acting on $\mathbb{R}^n$ since for all $s, t \in \mathbb{R}$ $\varsigma_s \circ \varsigma_t = \varsigma_{s+t}$, and ς_0 is the identity transformation. The group of translations is another abelian group action on $\mathbb{R}^n$, while the rotations are a non-abelian group of transformation on $\mathbb{R}^n$.

8 Lebesgue measure on the sphere S^{n-1}

We let S^{n-1} denote the unit sphere in $\mathbb{R}^n$; i.e., $S^{n-1} = \{x \in \mathbb{R}^n \ : \ \|x\| = 1 \ \}$. Define the map $\Omega : \mathbb{R}^n \backslash \{0\} \to S^{n-1}$ by

$$\Omega(x) = \frac{x}{\|x\|} \ .$$

Notice the Ω is continuous, and hence a Borel transformation.

8.1 DEFINITION (Lebesgue measure on S^{n-1})**.** Let X denote the punctured closed unit ball in $\mathbb{R}^n$; i.e, $X = \{x \in \mathbb{R}^n \ : \ 0 < \|x\| \le 1 \ \}$, and let μ denote the restriction of Lebesgue measure m_n to X. That is, for E a Borel set in $\mathbb{R}^n$,

$$\mu(E) = m_n(E \cap X) \ .$$

We define the *Lebesgue measure on* S^{n-1} to be the Borel measure σ_n on S^{n-1} given by

$$\sigma_n(A) = n\Omega\#\mu(A) = n\mu(\Omega^{-1}(A)) \tag{8.1}$$

for all Borel sets $A \subset S^n$.

8.2 Remark. The factor of n in the definition of σ_n is to make $\sigma_n(S^{n-1})$ equal to n times the Lebesgue measure of the unit ball which will be seen to be the "right" normalization below.

Notice that the map Ω commutes with rotations. That is, if $T : \mathbb{R}^n \to \mathbb{R}^n$ is any rotation

$$T\Omega(x) = \Omega T(x)$$

for all $x \in \mathbb{R}^n$ since $\|T(x)\| = \|x\|$ for all x, and T is linear. This has the following consequence:

8.3 THEOREM (Invariance of σ_n under rotations)**.** *Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be any rotation, which then acts on S^{n-1} be restriction.*

$$T\#\sigma_n = \sigma_n \ . \tag{8.2}$$

Proof. Let A be any Borel set in S^{n-1}, and T any rotation. Then, since T^{-1} is also a rotation, $\Omega^{-1} \circ T^{-1} = T^{-1} \circ \Omega^{-1}$. Therefore,

$$\sigma_n(T^{-1}A) = n\mu(\Omega^{-1}T^{-1}(A)) = n\mu(T^{-1}\Omega^{-1}(A)) = nT\#\mu(\Omega^{-1}(A)) = n\mu(\Omega^{-1}(A)) = \sigma_n(A) \ .$$

□

We now define a homeomorphism $\Phi : (0,\infty) \times S^n$ onto $\mathbb{R}^n\backslash\{0\}$ by

$$\Phi(r,\omega) = r\omega \ .$$

Since both Φ and Φ^{-1} are continuous, they are both Borel. Notice that

$$\Omega^{-1}(A) \cap X = \Phi((0,1] \times A) \ ,$$

and therefore

$$\frac{\sigma_n(A)}{n} = m_n(\Phi((0,1] \times A)) \ . \tag{8.3}$$

8.4 LEMMA. *For any $a < b \in (0,\infty)$, and any Borel set $A \subset S^{n-1}$,*

$$m_n(\Phi((a,b] \times A)) = \frac{b^n - a^n}{n}\sigma_n(A) \ . \tag{8.4}$$

Proof. Notice that (8.3) is the special case of (8.4) corresponding to $a = 0$ and $b = 1$. To get the general case, we use the dilation properties of Lebesgue measure. Let t be such that $e^{-t} = b$. Then

$$\Phi((0,b] \times A) = \varsigma_t(\Phi((0,1] \times A)) = \varsigma_{-t}^{-1}(\Phi((0,1] \times A)) \ ,$$

and therefore,

$$m_n(\Phi((0,b] \times A)) = \varsigma_{-t}\#m_n(\Phi((0,1] \times A)) = e^{-tn}m_n((0,1] \times A) \ ,$$

where in the last equality we have used (8.2). Combining this with (8.3) and recalling that $e^{-t} = b$, we obtain

$$m_n(\Phi((0,b] \times A)) = \frac{b^n}{n}\sigma_n(A) \ .$$

But then

$$m_n(\Phi((a,b] \times A)) = m_n(\Phi((0,b] \times A)) - m_n(\Phi((0,a] \times A))$$

and we obtain (8.4). □

8.5 Remark. Taking $a = 0$, $b = 1$ and $A = S^{n-1}$, we see that with B_n denoting the unit ball on $\mathbb{R}^n$,

$$m_n(B_n) = \frac{1}{n}\sigma_n(S^{n-1}) \ ,$$

and that with this normalization of σ_n,

$$\sigma_n(S^{n-1}) = \lim_{r\uparrow 1} \frac{m_n(\Phi((r,1] \times S^{n-1})}{1-r} \ ,$$

where $\Phi((r,1] \times S^{n-1})$ is a spherical shell of thickness $1-r$. Thus, the factor of n in (8.1) is natural.

8.6 DEFINITION. Let $F_n(t) = t^n/n$ for $t \in (0,\infty)$. Since F_n is right continuous, There is a unique Lebesgue-Stieltjes measure ϱ_n on $(0,\infty)$ such that for all $a < b \in (0,\infty)$,

$$\varrho_n((a,b]) = F_n(b) - F_n(a) \ .$$

We write ϱ_n to denote this measure in what follows.

8.7 Remark. Note that for $a < b \in (0,\infty)$,

$$\varrho_n((a,b]) = \int_a^b r^{n-1}\mathrm{d}r \ .$$

The point of this definition is that we may now rewrite (8.4) as

$$m_n(\Phi((a,b] \times A)) = \varrho_n \otimes \sigma_n((a,b] \times A) \ ,$$

and thus,

$$\Phi^{-1}\#m_n((a,b] \times A) = \varrho_n \otimes \sigma_n((a,b] \times A) \ , \tag{8.5}$$

Let $\mathcal{S}_n$ be the algebra consisting of all disjoint unions of sets of the form $(a,b] \times A$, $a < b \in (0,\infty)$ and $A \in \mathcal{B}_{S^{n-1}}$. Then $\sigma(\mathcal{S}_n)$ is easily seen to be the product σ-algebra $\mathcal{B}_{(0,\infty)} \otimes \mathcal{B}_{S^{n-1}}$. By the uniqueness theorem, we conclude that

$$\Phi^{-1}\#m_n = \varrho_n \otimes \sigma_n \qquad \text{and consequently} \qquad m_n = \Phi\#(\varrho_n \otimes \sigma_n) \ .$$

This identification of m_n as the push-forward under Φ of $\varrho_n \otimes \sigma_n$ leads to the following theorem for integration in polar coordinates:

8.8 THEOREM (Integration in polar coordinates)**.** *Let $f \in L^+(\mathbb{R}^n)$. Then*

$$\int_{\mathbb{R}^n} f\mathrm{d}m_n = \int_{(0,\infty)} \left[\int_{S^{n-1}} f(r\omega)\mathrm{d}\sigma_n(\omega)\right]\mathrm{d}\varrho_n(r) = \int_{S^{n-1}} \left[\int_{(0,\infty)} f(r\omega)\mathrm{d}\varrho_n(r)\right]\mathrm{d}\sigma_n(\omega) \ . \tag{8.6}$$

Proof. For $r \in (0,\infty)$ and $\omega \in S^{n-1}$, $\Phi(r,\omega) = r\omega$, and hence

$$\int_{(0,\infty)\times S^{n-1}} f(r\omega)\mathrm{d}\varrho_n \otimes \sigma_n = \int_{(0,\infty)\times S^{n-1}} f\circ\Phi(r,\omega)\mathrm{d}\varrho_n \otimes \sigma_n = \int_{\mathbb{R}^n\backslash\{0\}} f\mathrm{d}m_n = \int_{\mathbb{R}^n} f\mathrm{d}m_n \ .$$

Finally, by Tonelli's Theorem we have (8.6). □

In particular, since

$$\int_{\mathbb{R}} e^{-t^2/2}\mathrm{d}t = \sqrt{2\pi} \ ,$$

by Tonelli's Theorem once more,

$$\begin{aligned}
(2\pi)^{n/2} &= \int_{\mathbb{R}^n} e^- \|x\|^2/2\mathrm{d}m_n \\
&= \sigma_n(S^{n-1})\int_{(0,\infty)} e^{-r^2/2}r^{n-1}\mathrm{d}r \\
&= \sigma_n(S^{n-1})2^{n/2-1}\int_{(0,\infty)} e^{-u}u^{n/2-1}\mathrm{d}u \\
&= \sigma_n(S^{n-1})2^{n/2-1}\Gamma\left(\frac{n}{2}\right) \ .
\end{aligned}$$

Therefore,

$$\sigma_n(S^{n-1}) = \frac{2\pi^{n/2}}{\Gamma(n/2)} \quad \text{and} \quad m_n(B_n) = n\frac{2\pi^{n/2}}{\Gamma(n/2)} \ .$$

9 Properties of Lebesgue measure on S^{n-1}

There is an important identification of S^{n-1} with the one-point compactification of $\mathbb{R}^{n-1}$ through the *stereographic projection.*

9.1 DEFINITION (Stereographic projection)**.** Let $\omega_0 = -\mathbf{e}_n$ be the "South Pole" in S^{n-1}. For any other $\omega \in S^{n-1}$, define $T(\omega)$ to be the intersection of the line through ω_0 and ω with the hyperplane $\{x \in \mathbb{R}^n \ : \ \langle \mathbf{e}_n, x\rangle = 0\}$, which we may identify with $\mathbb{R}^{n-1}$ in the natural way. Then $T : S^{n-1}\backslash\omega_0 \to \mathbb{R}^{n-1}$ is the stereographic of $S^{n-1}\backslash\omega_0$ onto $\mathbb{R}^{n-1}$. Let $\mathbb{R}^{n-1}\cup\infty$ be the one point compactification of $\mathbb{R}^{n-1}$, so that the neighborhoods of ∞ are the complements of compact sets in $\mathbb{R}^{n-1}$. Then we define $T(\omega_0) = \infty$, which yields the stereographic projection of S^{n-1} onto the one-point compactification of $\mathbb{R}^{n-1}$.

It is easy to work out a formula for T. Let us write vectors in $\mathbb{R}^n$ in the form (v,z) where $\mathbf{v} \in \mathbb{R}^{n-1}$ and $z \in \mathbb{R}$. Then for any $\omega = (v,z)$ other than ω_0, the line through ω_0 and ω is parameterized by

$$(1-t)\omega_0 + t\omega = (tv, t(1+z)-1) \ .$$

Then $\langle (tv, t(1+z)-1), \mathbf{e}_n\rangle = 0$ reduces to $t = (1+z)^{-1}$, and hence

$$T((v,z)) = \frac{1}{1+z}v \ . \tag{9.1}$$

It is also useful to have a formula for the inverse. Let $x := T((v, z)$, so that $(1 + z)x = v$. Since $(v, z) \in S^{n-1}$, $\|v\|^2 + z^2 = 1$, and so $(1+z)^2\|x\|^2 = 1 - z^2$ which is readily solved for z in terms of $\|x\|^2$, and then $(1 + z)x = v$ gives us v:

$$T^{-1}(x) = \frac{1}{1 + \|x\|^2}(2x, 1 - \|x\|^2) \ . \tag{9.2}$$

The map T is evidently a homeomorphism of $S^{n-1}\backslash\omega_0$ onto $\mathbb{R}^{n-1}$. We may use it to transfer the half open rectangle algebra of sets $\mathbb{R}^{n-1}$ to an algebra of sets in $S^{n-1}\backslash\omega_0$, since any bijective image of an algebra is an algebra. Call this algebra $\mathcal{A}_{S^{n-1}}$. We know that every open set in $\mathbb{R}^{n-1}$ can be written as a countable union of sets in the half-open rectangle algebra on $\mathbb{R}^{n-1}$, and then, since T is a homeomorphism, every open set in $S^{n-1}\backslash\omega_0$ is a countable union of sets in $\mathcal{A}_{S^{n-1}}$. It follows that

$$\mathcal{B}_{S^{n-1}} = \sigma(\mathcal{A}_{S^{n-1}}) \ .$$

Now let μ be *any* Borel measure μ on $S^{n-1}\backslash\omega_0$ such that $\mu(S^{n-1}) < \infty$. By our general results concerning measures on σ-algebras generated by algebras, it follows that every set $E \in \mathcal{B}_{S^{n-1}\backslash\omega_0}$ has the property that for every $\epsilon > 0$, there is a set $A \in \mathcal{A}_{S^{n-1}}$ such that

$$\mu(E\Delta A) \le \epsilon \ . \tag{9.3}$$

(The condition that $\mu(S^{n-1}) < \infty$ ensures that $\mu(E) < \infty$, a requirement of the general theorem.)

It then follows, in the usual way, that every $f \in L^1(S^{n-1}\backslash\omega_0, \mathcal{B}_{S^{n-1}\backslash\omega_0}, \mu)$ may be approximated by a really simple function, and then rounding the corners", by a continuous function.

Finally, if $\mu(\{\omega_0\}) = 0$, there is no difference between $L^1(S^{n-1}\backslash\omega_0, \mathcal{B}_{S^{n-1}\backslash\omega_0}, \mu)$ and $L^1(S^{n-1}, \mathcal{B}_{S^{n-1}}, \mu)$. We have proved:

9.2 THEOREM. *Let μ be any Borel measure on S^{n-1} such that $\mu(S^{n-1}) < \infty$ and such that $\mu(\omega) = 0$ for all single points ω. Then $L^1(S^{n-1}, \mathcal{B}_{S^{n-1}}, \sigma_n)$ is separable, and the continuous functions $C(S^{n-1})$ are dense in it.*

We next prove a uniqueness theorem for Lebesgue measure on S^{n-1}.

9.3 THEOREM. *Let μ be any Borel measure on S^{n-1} such that $\mu(S^{n-1}) < \infty$, and such that for all rotations R,*

$$R\#\mu = \mu \ .$$

Then

$$\sigma_n(S^{n-1})\mu = \mu(S^{n-1})\sigma_n \ .$$

The proof of this theorem is based on some lemmas of independent interest that we now present. In what follows, L^2 denotes $L^2(S^{n-1}, \mathcal{B}_{S^{n-1}}, \sigma_n)$ and $\|f\|_2^2 = \int_{S^{n-1}} |f|^2 \mathrm{d}\sigma_n$.

Given any $f \in L^2$, we define Rf to be the function $f \circ R$. Then $f \to Rf$ is linear, and since σ_n is rotation invariant,

$$\|f\|_2 = \|Rf\|_2 \ .$$

Next, for any rotation R define

$$A_R f = \frac{1}{2}(f + R) \ .$$

This operation is also linear, and it is a contraction in the L^2 norm by the Minkowski inequality and the fact that $\|Rf\|_2 = \|f\|_2$. Moreover, if f is continuous, Rf has the same modulus of continuity as f, and the average of f and Rf, namely $A_R f$, therefore has a modulus of continuity no greater than that of f.

9.4 LEMMA. *For any real valued $f \in L^2$, and any rotation R,*

$$\|A_R f\|_2 \le \|f\|_2 \ ,$$

and there is equality if and only if $Rf = f$ in L^2. More generally, for any finite set $\{R_1, \ldots, R_m\}$ of rotations,

$$\|A_{R_m} \cdots A_{R_1} f\|_2 \le \|f\|_2 \ ,$$

and there is equality if and only if $R_j f = f$ for each $j = 1, \ldots, m$.

Proof. We compute

$$\left\|\frac{1}{2}(f + Rf)\right\|_2^2 = \frac{1}{4}(\|f\|_2^2 + \|Rf\|_2^2 + 2\int_{S^{n-1}} fRf \mathrm{d}\sigma_n \ .$$

However, by the rotation invariance of σ_n, $\|Rf\|_2 = \|f\|_2$, and by the Cauchy-Schwarz inequality,

$$\int_{S^{n-1}} fRf \mathrm{d}\sigma_n \le \|f\|_2 \|Rf\|_2 = \|f\|_2^2 \ .$$

This proves the inequality. Note that there is equality in the Cauchy-Schwarz inequality if and only if $\|Rf\|_2 f = \|f\|_2 Rf$, which reduces to $f = Rf$.

For the second part, consider first the case $m = 2$, and suppose $\|A_{R_2} A_{R_1} f\|_2 = \|f_2\|$. By the first part,

$$\|A_{R_2} A_{R_1} f\|_2 \le \|A_{R_1} f\|_2 \qquad \text{and} \qquad \|A_{R_1} f\|_2 \le \|f\|_2 \ .$$

We must have equality in both inequalities. By what we have proved above, equality on the right implies that $R_1 f = f$, and then of course $A_{R_1} f = f$. Then the inequality on the left reduces to $\|A_{R_2} f\|_2 \le \|f\|_2$, and be what we have proved above, equality here implies that $R_2 f = f$. The general case follows in the same way. □

9.5 LEMMA. *Let f be any continuous functions on S^{n-1}. Let $\{R_1, \ldots, R_m\}$ be any finite set of rotations. Define a sequence $\{f_j\}$ by*

$$f_0 = f \qquad \text{and} \qquad f_{j+1} = A_{R_m} \cdots A_{R_1} f_j \ .$$

Then f_j converges uniformly to a continuous function h such that

$$R_j h = h$$

for $j = 1, \ldots, m$. Moreover, there is a choice of $m = n - 1$ rotations $\{R_1, \ldots, R_{n-1}\}$ for which the limiting function h is constant.

Proof. By what we have noted above, the sequence $\{f_j\}$ is uniformly equicontinuous and equibounded, by Arzela-Ascoli Theorem, there is a subsequence $\{f_{j_k}\}$ and $h \in C(S^{n-1})$ such that $f_{j_k} \to h$ uniformly.

Next note that by the lemma, $\|f_j\|_2$ is monotone decreasing. Define

$$c = \lim_{j\to\infty} \|f_j\|_2 \ .$$

Since uniform convergence implies L^2 convergence for finite measure spaces,

$$\|h\|_2 = \lim_{k\to\infty} \|f_{j_k}\|_2 = c \ .$$

Since the linear operator $A_{R_m} \cdots A_{R_1}$ is continuous (it is even a contraction),

$$[A_{R_m} \cdots A_{R_1}] f_{j_k} \to [A_{R_m} \cdots A_{R_1}] h$$

uniformly. But the left hand side is f_{j_k+1}, and so

$$\|A_{R_m} \cdots A_{R_1} h\|_2 = \lim_{k\to\infty} \|f_{j_k+1}\|_2 = c \ .$$

That is,

$$\|A_{R_m} \cdots A_{R_1} h\|_2 = \|h\|_2 \ .$$

The lemma now implies that h is invariant under each of $R_1, \ldots, B_m$.

Let $\|\cdot\|_\infty$ denote the supremum norm, which gives the uniform topology. Since for any continuous g,

$$\|A_{R_m} \cdots A_{R_1} g - h\|_\infty = \|A_{R_m} \cdots A_{R_1}(g-h)\|_\infty \le \|g-h\|_\infty \ ,$$

The fact that $\|f_{j_k} - h\|_\infty \to 0$ implies that $\|f_j - h\|_\infty \to 0$ along the whole sequence. The final part is left as an exercise for the reader. □

Proof of Theorem 9.3. We may suppose $\mu(S^{n-1}) \neq 0$, or else the claim is trivial. Then normalizing, we may suppose without loss of generality that $\mu(S^{n-1}) = \sigma_n(S^{n-1})$.

Next, if $\mu(\{\omega\}) = c > 0$ for some ω, $\mu(\{R\omega\}) = c$ for every rotation R. We can choose an infinite sequence of rotations to obtain an infinite sequences of distinct points in this way. This would force $\mu(S^{n-1}) = \infty$, and so μ does not charge single points. Thus, Theorem 9.2 applies to μ, σ_n and to $\mu + \sigma_n$.

It suffices to show that

$$\int_{S^{n-1}} f \mathrm{d}\mu = \int_{S^{n-1}} f \mathrm{d}\sigma_n \tag{9.4}$$

for all $f \in C(S^{n-1})$. This is because if E is any Borel set, 1_E may be approximated by continuous functions in $L^1(S^{n-1}, \mathcal{B}_{S^{n-1}}, \mu + \sigma_n)$ by Theorem 9.2. But if

$$\lim_{n\to\infty} \|f_n - 1_E\|_{L^1(S^{n-1}, \mathcal{B}_{S^{n-1}}, \mu+\sigma_n)} = 0 \ ,$$

then $f_n \to 1_E$ in both $L^1(S^{n-1}, \mathcal{B}_{S^{n-1}}, \mu)$ and $L^1(S^{n-1}, \mathcal{B}_{S^{n-1}}, \sigma_n)$.

Therefore, if (9.4) is true for all continuous functions f,

$$\begin{aligned}\mu(E) = \int_{S^{n-1}} 1_E \mathrm{d}\mu &= \lim_{n\to\infty} \int_{S^{n-1}} f_n \mathrm{d}\mu \\ &= \lim_{n\to\infty} \int_{S^{n-1}} f_n \mathrm{d}\sigma_n = \int_{S^{n-1}} 1_E \mathrm{d}\sigma_n = \sigma_n(E) \ .\end{aligned}$$

Next, given any continuous function f, in the lemmas above we have constructed a sequence $\{f_j\}$ of continuous function that converges uniformly to constant functions h. Moreover since each f_j is an average over rotations of f and since both μ and σ_n are rotation invariant,

$$\int_{S^{n-1}} f_j \mathrm{d}\sigma_n = \int_{S^{n-1}} f \mathrm{d}\sigma_n \quad \text{and} \qquad \int_{S^{n-1}} f_j \mathrm{d}\mu = \int_{S^{n-1}} f \mathrm{d}\mu$$

for all j .

Since uniform convergence implies convergence of integrals on a finite measure space,

$$\int_{S^{n-1}} f_j \mathrm{d}\sigma_n \to \int_{S^{n-1}} h \mathrm{d}\sigma_n \quad \text{and} \qquad \int_{S^{n-1}} f_j \mathrm{d}\mu \to \int_{S^{n-1}} h \mathrm{d}\mu$$

as $j \to \infty$. But since h is constant and μ and σ_n have the same total mass,

$$\int_{S^{n-1}} h \mathrm{d}\sigma_n = \int_{S^{n-1}} h \mathrm{d}\mu \ .$$

Combining the last three identities, we obtain (9.4) $\square$