CHAPTER 1

Introduction to Topological Groups and the Birkhoff-Kakutani Theorem

1. Introduction

For us, a topological group is a group G that is equipped with a topology that makes the functions $(x, y) \mapsto xy$ from $G \times G$ to G and $x \mapsto x^{-1}$ from G to G continuous.

Here are some basic observations regarding topological groups; they follow simply and directly from the definition

- given $a, b \in G$, should ab find itself in an open set U then there are open sets V and W such that $a \in V, b \in W$ and $V \cdot W = \{xy : x \in V, y \in W\} \subseteq U$.
- given $a \in G$, should U be an open set containing a^{-1}, then there is an open set V containing a so that $V^{-1} = \{v^{-1} : v \in V\} \subseteq U$.
- given $a, b \in G$, should U be an open set containing ab^{-1}, then there are open sets V and W such that $a \in V, b \in W$ and $V \cdot W^{-1} \subseteq U$.
- given $a, b \in G$, should U be an open set containing $a^{-1}b$, then there are open sets V and W such that $a \in V, b \in W$ and $V^{-1} \cdot W \subseteq U$.
- Each of the mappings from G to G

$$\begin{aligned} l_a &: G \to G, & l_a(x) &= ax \\ r_a &: G \to G, & r_a(x) &= xa \\ \mathrm{inv} &: G \to G, & \mathrm{inv}(x) &= x^{-1} \end{aligned}$$

  is a homeomorphism of G onto G.
- If F is a closed subset of G then so are aF, Fa, and F^{-1} for any $a \in G$.
- If U is an open subset of G and S is a non-void subset of G then the sets $S \cdot U, U \cdot S$, and U^{-1} are open subsets of G.
- G **is homogeneous**: if $p, q \in G$ then there is a homeomorphism ϕ of G onto G such that $\phi(p) = q$. Indeed $\phi = l_{qp^{-1}}$ will do the trick.

The homogeneity of topological groups has consequences regarding its topological structure. Here's one such

Theorem 1.0.1. *Let G be a topological group. If U is an open set containing the identity e then there is an open set V containing e such that $e \in V \subseteq \overline{V} \subseteq U$. Consequently, a T_0 topological group is regular and so Hausdorff.*

Proof. First things first: Let U be an open set that contains the identity e. By continuity of multiplication there is an open set W containing e such that $W \cdot W \subseteq U$. If we set $V = W \cap W^{-1}$ then we have an open set that contains e, is symmetric ($V = V^{-1}$) and satisfies $V \cdot V \subseteq U$.

We claim $\overline{V} \subseteq U$. Take $x \in \overline{V}$. Then xV is an open set that contains x and so $xV \cap V \neq \emptyset$, so there must be $v_1, v_2 \in V$ so that $xv_1 = v_2$. But then

$$x = v_2 v_1^{-1} \in V \cdot V^{-1} = V \cdot V \subseteq U.$$

The homogeneous structure of G now tells us that whenever $x \in G$ and U is an open set containing x, then there is an open set V such that $x \in V \subseteq \overline{V} \subseteq U$. Regularity follows from this and G's homogeneity. Again, homogeneity of G tells us that if $x, y \in G$ are distinct and there's a neighborhood of x that doesn't contain y, then we have a neighborhood of y that doesn't contain x. In other words, T_0 topological groups satisfy the T_1 axiom. □

Henceforth, we assume that all topological groups are Hausdorff.

The surprising conclusion reached in Theorem 1.0.1 is a typical product of the mix of the algebra and topology in topological groups. Here's another:

Proposition 1.0.2. *Every open subgroup of a topological group is closed.*

Proof. Let H be an open subgroup of the topological group G. Take $g \in \overline{H}$. Every open set that contains g intersects H; gH is such an open set. Therefore $gH \cap H \neq \emptyset$. Since cosets are either the same or disjoint, $gH = H$. Thus

$$g = ge \in gH = H,$$

and $\overline{H} \subseteq H$. □

Exercise 1. *If G is a connected topological group then any neighborhood of the identity is a system of generators for G.*

2. The Classical (locally compact) groups

$\mathbb{R}^n$ and $\mathbb{C}^n$ will denote, as usual, real and complex $n-$spaces, respectively. M_n will denote the linear algebra of all $n \times n$ matrices with complex entries. We can associate with any $(a_{ij}) \in M_n$ the point $(b_1, \dots, b_{n^2}) \in \mathbb{C}^{n^2}$, where

$$b_{i+(j-1)n} = a_{ij};$$

this establishes a bijective correspondence between M_n and $\mathbb{C}^{n^2}$, a correspondence we use to equip M_n with the Euclidean topology of $\mathbb{C}^{n^2}$.

Of course if $\alpha, \beta \in M_n$ then $\alpha \cdot \beta \in M_n$, too, with

$$c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}$$

whenever $\alpha = (a_{ij}), \beta = (b_{ij})$, and $\alpha \cdot \beta = (c_{ij})$. It is easy to see that the operation $(\alpha, \beta) \to \alpha \cdot \beta$ is continuous from $M_n \times M_n$ to M_n.

In addition to addition and multiplication, M_n is endowed by nature with a couple of other natural operations: transposition and conjugation: if $\alpha = (a_{ij}) \in M_n$ then $\alpha^t = (a_{ji})$ and $\bar{\alpha} = (\bar{a}_{ij})$, where $\bar{a}$ is the complex conjugate of the complex number a. Both of these operations are homeomorphisms of M_n onto itself and each is of 'order two', that is, $\alpha^{tt} = \alpha$ and $\bar{\bar{\alpha}} = \alpha$.

Some members of M_n have a multiplicative inverse; the collection of all such matrices will be denoted by $GL(n,\mathbb{C})$ and called the **General Linear Group.** Of course, $\alpha \in M_n$ belongs to $GL(n,\mathbb{C})$ precisely when $\det\alpha \neq 0$. Now it is easy to believe and also true that $\alpha \to \det\alpha$ is a continuous function of the coordinates of α and so $GL(n,\mathbb{C})$ is an open subset of M_n. If $\alpha,\beta \in GL(n,\mathbb{C})$ then $\alpha\cdot\beta \in GL(n,\mathbb{C})$ and $(\alpha\cdot\beta)^{-1} = \beta^{-1}\cdot\alpha^{-1}$. This is elementary linear algebra; further if $\alpha \in GL(n,\mathbb{C})$ then α^{-1} has coordinates b_{ij} where

$$b_{ij} = \frac{p_{ij}(\alpha)}{\det(\alpha)}$$

where $p_{ij}(\alpha)$ is a polynomial with coordinates of α. It follows that the operation $\alpha \to \alpha^{-1}$ of $GL(n,\mathbb{C})$ onto itself is also a homeomorphism (of order two).

COROLLARY 2.0.3. *$GL(n,\mathbb{C})$ is a locally compact metrizable topological group.*

PROOF. After all,

$$GL(n,\mathbb{C}) = \det^{\leftarrow}(\{z \in \mathbb{C} : z \neq 0\}),$$

and so $GL(n,\mathbb{C})$ is homeomorphic to an open subset of a locally compact metric space, $\mathbb{C}^{n^2}$. Our comments about continuity of the operations $(\alpha,\beta) \to \alpha\cdot\beta$ and $\alpha \to \alpha^{-1}$ finish the proof. □

Inside $GL(n,\mathbb{C})$ we can find other *classical* topological groups. Here are a few of them.
$\mathbf{O(n)}$: $\alpha \in GL(n,\mathbb{C})$ is **orthogonal** if $\alpha = \bar{\alpha} = (\alpha^{-1})^t$.
$\mathbf{O(n,\mathbb{C})}$: $\alpha \in GL(n,\mathbb{C})$ is **complex orthogonal** if $\alpha = (\alpha^{-1})^t$.
$\mathbf{U(n)}$: $\alpha \in GL(n,\mathbb{C})$ is **unitary** if $\bar{\alpha} = (\alpha^{-1})^t$.

Since the mappings $\alpha \to \alpha^{-1}$ and $\alpha \to (\alpha^{-1})^t$ are continuous in $GL(n,\mathbb{C})$, **each of the groups** $O(n), O(n,\mathbb{C})$ **and** $U(n)$ **are closed subgroups of** $GL(n,\mathbb{C})$.

$\mathbf{GL(n,\mathbb{R})}$: $\alpha \in M_n$ is real if $\alpha = \bar{\alpha}$; denote the set of real members of M_n by $M_n(\mathbb{R})$ so

$$GL(n,\mathbb{R}) = GL(n,\mathbb{C}) \cap M_n(\mathbb{R}).$$

$\mathbf{SL(n,\mathbb{C})}$: the members of $GL(n,\mathbb{C})$ with determinant 1, called the **Special Linear Group.**
$\mathbf{SL(n,\mathbb{R})}$: $SL(N,\mathbb{C}) \cap M_n(\mathbb{R})$.
$\mathbf{SO(n)}$: $SL(n,\mathbb{C}) \cap O(n)$.
$\mathbf{SU(n)}$: $SL(n,\mathbb{C}) \cap U(n)$.
Again, $SL(n,\mathbb{C}), SL(n,\mathbb{R}), SO(n), SU(n)$ **are closed subgroups of** $GL(n,\mathbb{C})$). Actually more can be said.

THEOREM 2.0.4. *The groups $U(n), O(n), SU(n)$ and $SO(n)$ are compact metric topological groups.*

PROOF. Each of $O(n), SU(n), SO(n)$ are closed subgroups of $U(n)$, so it's enough to establish that $U(n)$ is compact. Now $\alpha \in U(n)$ precisely when $\alpha^t\bar{\alpha} = \mathrm{id}|_{\mathbb{C}^n}$. This in turn is the same as saying if $\alpha = (a_{ij})$ that for each $1 \le i,k \le n$,

$$\sum_j a_{ji}\bar{a}_{jk} = \delta_{ik}.$$

Now the left side is a continuous function of α so $U(n)$ is closed not just in $GL(n,\mathbb{C})$ but even in M_n. Moreover

$$\sum_j a_{ji}\bar{a}_{ji} = 1$$

February 3, 2009

ensures that $|a_{ij}| \leq 1$ for $1 \leq i, j \leq n$. So the entries of any $\alpha \in U(n)$ are bounded. But this ensures that $U(n)$ is homeomorphic to a closed bounded subset of $\mathbb{C}^{n^2}$. □

3. The Birkhoff-Kakutani Theorem

Our next result is technical but it's an investment well-worth the price.

THEOREM 3.0.5 (Garrett Birkhoff/Shizuo Kakutani). *Let (U_n) be a sequence of symmetric open sets each containing the identity e of the topological group G. Suppose that for each $k \in \mathbb{N}$,*

$$U_{k+1} \cdot U_{k+1} \subseteq U_k.$$

Let $H = \cap_k U_k$. Then there is a left invariant pseudo-metric σ on $G \times G$ such that

(i) *σ is 'left uniformly continuous' on $G \times G$;*
(ii) *$\sigma(x, y) = 0$ is and only if $x \in yH$;*
(iii) *$\sigma(x, y) \leq 4 \cdot 2^{-k}$, if $x \in y \cdot U_k$;*
(iv) *$2^{-k} \leq \sigma(x, y)$, if $x \notin yU_k$.*

PROOF. We start by reinterpreting the sequence's descending character with an eye toward defining σ. For each k, set

$$V_{2^{-k}} = U_k.$$

Next define V_r for r a dyadic rational number with $0 < r < 1$ as follows: if

$$r = 2^{-l_1} + 2^{-l_2} + \cdots + 2^{-l_n},$$

with $0 < l_1 < \cdots < l_n$, all positive integers, let

$$V_r = V_{2^{-l_1}} \cdot V_{2^{-l_2}} \cdots V_{2^{-l_n}}.$$

For dyadic r's greater than or equal to 1, set $V_r = G$. Here's what's so

$$r < s \Rightarrow V_r \subseteq V_s; \tag{3.1}$$

further for any $l \in \mathbb{N}$, we have

$$V_r \cdot V_{2^{-l}} \subseteq V_{r+2^{-l+2}}. \tag{3.2}$$

We put off the somewhat tedious yet clever proofs of the indicated relationships between the V's (equations (3.1) and (3.2)) until the end of our general discussion.

With these in hand we go forth to define σ. First, for $x \in G$, let

$$\phi(x) = \inf\{r : x \in V_r\}.$$

Plainly $\phi(x) = 0$ if and only if $x \in H$. Now for $x, y \in G$ define

$$\sigma(x, y) = \sup\{|\phi(zx) - \phi(zy)| : z \in G\}.$$

Plainly, $\sigma(x, y) = \sigma(y, x)$ and $\sigma(x, x) = 0$. It's easy to see that $\sigma(x, u) \leq \sigma(x, y) + \sigma(y, u)$ for any $x, y, u \in G$, and the fact that $\sigma(ax, ay) = \sigma(x, y)$ for all $a \in G$ is obvious. So σ is a left invariant pseudometric.

Now we join the hunt.

Let $l \in \mathbb{N}$, and suppose $u \in V_{2^{-l}}$ and $z \in G$. If $z \in V_r$ then $z \cdot u \in V_{r+2^{-l+2}}$ thanks to (3.2). Hence

$$\phi(z \cdot u) \leq r + 2^{-l+2};$$

this is true whenever $z \in V_r$, so using the definition of ϕ,

$$\phi(z \cdot u) \leq \phi(z) + 2^{-l+2}.$$

Similarly, if $z \cdot u \in V_r$ then

$$z \in V_r \cdot u^{-1} \subseteq V_r V_{2^{-l}}^{-1} = V_r V_{2^{-l}} \subseteq V_{r+2^{-l+2}},$$

again by (3.2). It follows that

$$\phi(z) \leq r + 2^{-l+2},$$

so we see that

$$\phi(z) \leq \phi(z \cdot u) + 2^{-l+2}.$$

So

$$\phi(z) \leq \phi(z \cdot u) + 2^{-l+2} \quad \text{and} \quad \phi(z \cdot u) \leq \phi(z) + 2^{-l+2}.$$

The only conclusion that we can make is that for $u \in V_{2^{-l}}$, and $z \in G$,

$$|\phi(z) - \phi(z \cdot u)| \leq 2^{-l+2}.$$

From this we see that

$$\sigma(u, e) \leq 2^{-l+2}, \ \text{ for } \ u \in V_{2^{-l}}.$$

The third statement of the theorem follows from this and σ's left invariance: If $x \in yU_k$ then $y^{-1}x \in U_k = V_{2^{-k}}$ so

$$\sigma(x, y) = \sigma(y^{-1}x, e) \leq 2^{-k+2} = 4 \cdot 2^{-k}.$$

Next we deal with σ's 'uniform continuity.' Suppose $x, y, \tilde{x}$, and $\tilde{y}$ satisfy

$$y^{-1}x \in V_{2^{-l-1}} \quad \text{and} \quad \tilde{y}^{-1}\tilde{x} \in V_{2^{-l-1}}.$$

Then

$$\tilde{x}^{-1}\tilde{y}y^{-1}x \in V_{2^{-l-1}}^{-1} \cdot V_{2^{-l-1}} = V_{2^{-l-1}} \cdot V_{2^{-l-1}} \subseteq V_{2^{-l}},$$

and so

$$\begin{aligned} |\sigma(x, y) - \sigma(\tilde{x}, \tilde{y})| &= |\sigma(y^{-1}x, e) - \sigma(\tilde{y}^{-1}\tilde{x}, e)| \\ &\leq |\sigma(y^{-1}x, \tilde{y}^{-1}\tilde{x})| \\ &= |\sigma(\tilde{x}^{-1}\tilde{y}y^{-1}x, e)| \\ &\leq 2^{-l+2}, \ \text{(by (iii))} \end{aligned}$$

and this is what we mean by σ is 'left uniformly continuous.'

For (iv), suppose $y^{-1}x \notin U_l = V_{2^{-l}}$. Then $\phi(y^{-1}x) \geq 2^{-l}$ and so

$$\sigma(x, y) = \sigma(y^{-1}x, e) \geq \phi(y^{-1}x) \geq 2^{-l},$$

where the last inequality follows since for any $a \in G$,

$$\begin{aligned} \sigma(a,e) &= \sup\{|\phi(za)-\phi(z)| : z \in G\} \\ &\geq |\phi(a)-\phi(e)| = |\phi(a)|, \ \text{(since } e \in V_r \text{ for every } r.) \end{aligned}$$

Finally, (ii) is an easy consequence of (iii) and (iv). □

The hard work of Birkhoff and Kakutani pays off in a couple of fundamental consequences, consequences which underscore the special character of topological groups.

COROLLARY 3.0.6. *Let G be a topological group. If G has a countable neighborhood base at $\{e\}$ then G is metrizable. In this case the metric can be taken to be left invariant.*

PROOF. Suppose $\{V_n : n \in \mathbb{N}\}$ is a countable open base at e. Let $U_1 = V_1 \cap V_1^{-1}$, and U_2 be a symmetric open neighborhood of e such that such that $U_2 \subseteq U_1 \cap V_2$ and $U_2 \cdot U_2 \subseteq U_1$. Continuing, let U_n be a symmetric open neighborhood of e such that

$$U_n \subseteq U_1 \cap U_2 \cap \cdots \cap U_{n-1} \cap V_n,$$

and $U_n^2 \subseteq U_{n-1}$.

The family $\{U_k : k \in \mathbb{N}\}$ satisfies the conditions set forth in the Birkhoff-Kakutani theorem. Further, $H = \cap_n U_n = \{e\}$. Let σ be the left pseudo-metric introduced in the theorem. In fact, σ is a true metric; $\sigma(x,y) = 0$ if and only if $x = y$, since after all, $H = \{e\}$!

Since

$$\{x \in G : \sigma(x,e) < 2^{-k}\} \subseteq U_k \subseteq \{x \in G : \sigma(x,e) \leq 2^{-k+2}\},$$

the topology defined by σ coincides with the given topology of G. □

COROLLARY 3.0.7. *Let G be a topological group, $a \in G$, and F be a closed subset of G such that $a \notin F$. Then there is a continuous real function χ on G such that $\chi(a) = 0$ and $\chi(x) = 1$ for all $x \in F$. Consequently, every topological group is completely regular.*

PROOF. Let U_1 be a symmetric neighborhood of e such that $(aU_1) \cap F = \emptyset$. Choose a sequence $(U_n : n \geq 2)$ of open neighborhoods of e such that each U_n is symmetric, $U_{n+1} \cdot U_{n+1} \subseteq U_n$, and let $H = \cap_n U_n$. Apply Birkhoff-Kakutani theorem to the U_n's. For $x \in G$, define $\chi(x)$ by

$$\chi(x) = \min\{1, 2\sigma(a,x)\}$$

where σ is the left-invariant, uniformly continuous pseudometric produced in the Birkhoff-Kakutani theorem. χ is continuous by (i) of our theorem, and $\chi(a) = 0$. If $x \in F$ then $a^{-1}x \in a^{-1}F$, a set disjoint from U_1; consequently, $a^{-1}x \notin U_1$ and $\sigma(a,x) \geq 2^{-1}$ by (iv) of our theorem. It follows that $\chi(x) = 1$ for all $x \in F$. □

COROLLARY 3.0.8. *Let G be a locally compact group and suppose $\{e\}$ is the intersection of countably many open sets. Then G admits a left-invariant metric compatible with its topology.*

PROOF. Suppose $\{e\} = \cap_n U_n$, where U_n is open in G. Choose open sets $V_1, V_2, \ldots, V_n, \ldots$, each containing e with $\overline{V_n}$ a compact subset of $V_{n-1} \cap U_n$ for each n. We claim that $\{V_n : n \in \mathbb{N}\}$ is an open base at e.

Let W be an open set in G that contains e. Could it be that no $V_n \subseteq W$? Well let's suppose this to be the case. Then $\{\overline{V}_n \cap W^c : n \geq 2\}$ enjoys the finite intersection property; indeed,

$$\overline{V}_1 \cap \ldots \cap \overline{V}_n \cap W^c \supseteq \overline{V}_{n+1} \cap W^c \supseteq V_{n+1} \cap W^c \neq \emptyset.$$

It follows that $\cap_n \overline{V}_n \cap W^c \neq \emptyset$. But

$$\bigcap_{n\geq 2} (\overline{V}_n \cap W^c) = \left(\bigcap_{n\geq 2} \overline{V}_n\right) \cap W^c \subseteq \left(\bigcap_n U_n\right) \cap W^c = \{e\} \cap W^c = \emptyset.$$

OOPS! There must be some V_n that's contained in W. □

Here's another feature of topological groups that distinguishes them from general topological spaces; they admit apt notions of uniform continuity, for instance.

THEOREM 3.0.9. *Let G be a topological group and M be a non-empty compact subset of G. Then any continuous function $f : G \to \mathbb{R}$ is left uniformly continuous on M. i.e., given $\epsilon > 0$ there is an open set V containing the identity of G so that if $x, y \in M$ and $x \in yV$ then $|f(x) - f(y)| \leq \epsilon$.*

PROOF. Let $\epsilon > 0$ be given. For each $a \in M$ there is an open set V_a that contains the identity such that if $x \in M$ and $x \in aV_a$ then $|f(x) - f(a)| \leq \frac{\epsilon}{2}$. Since $e \cdot e = e$, there is an open set W_a that contains the identity e and satisfies $W_a \cdot W_a \subseteq V_a$. Now if $a \in M$ then $a \in a \cdot W_a$ so $\{a \cdot W_a : a \in M\}$ covers M; we can find $a_1, \ldots, a_n \in M$ so that

$$M \subseteq (a_1 \cdot W_{a_1}) \cup \cdots \cup (a_n \cdot W_{a_n}).$$

Look at

$$V = W_{a_1} \cap \cdots \cap W_{a_n}.$$

Then V is open and contains e. Let $x, y \in M$ with $x \in yV$. To see that $|f(x) - f(y)| \leq \epsilon$ note that if $y \in M$ then $y \in a_i \cdot W_{a_i}$ for some $i = 1, \ldots, n$. It follows from this and $W_a \subseteq V_a$ that $|f(y) - f(a_i)| \leq \frac{\epsilon}{2}$. Also

$$x \in yV \subseteq a_i \cdot W_{a_i} \cdot V \subseteq a_i \cdot W_{a_i} \cdot W_{a_i} \subseteq a_i \cdot V_{a_i},$$

and so $|f(x) - f(a_i)| \leq \frac{\epsilon}{2}$ too.

In sum, if $x, y \in M$ with $x \in yV$ then

$$|f(x) - f(y)| \leq |f(x) - f(a_i)| + |f(a_i) - f(y)| \leq \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.$$

□

We'll be spending much of our time in a locally compact setting and again, there's more than meets the eye because of the group's structure.

THEOREM 3.0.10. *Any locally compact topological group is paracompact, hence normal.*

PROOF. Let G be a locally compact topological group, and let V be an open set in G containing the identity of G and having a compact closure $\overline{V}$. By thereom 1.0.1 there exists a symmetric open set U such that

$$e \in U \subseteq \overline{U} \subseteq U \cdot U \subseteq V.$$

Look at $H = \cup_n U^n$, where $U^n = \underbrace{U \cdots U}_{n}$. Then H is an open subgroup (hence closed) of G. Also H is σ-compact since

$$H = \bigcup_n \underbrace{U \cdots U}_{2n} \supseteq \bigcup_n \underbrace{\overline{U} \cdots \overline{U}}_{n} \supseteq H,$$

a countable union of compact subsets. Hence H is Lindelöf and regular, so H is paracompact [**?**].

Let $\mathcal{U}$ be an open cover of G. Each coset xH of H is also σ-compact and so there is a countable subfamily $\{V_{xH}^{(n)} : n \in \mathbb{N}\}$ of $\mathcal{U}$ that covers xH.

Naturally,

$$\{V_{xH}^{(n)} \cap xH : n \in \mathbb{N}\}$$

is an open cover of xH and since xH is paracompact, there is a locally finite open cover of xH that refines $\{V_{xH}^{(n)} \cap xH : n \in \mathbb{N}\}$, call it $\{\mathcal{W}_{xH}^{(n)}\}_{n=1}^{\infty}$. Note that for each $n = 1, 2, \ldots$

$$\mathcal{W}_{xH}^{(n)} \subseteq xH.$$

For each $n = 1, 2 \ldots$, let

$$\mathcal{W}^n = \bigcup_{xH \in G/H} \mathcal{W}_{xH}^{(n)},$$

so $\{\mathcal{W}^n\}_{n=1}^{\infty}$ is a locally finite open cover. Therefore $\mathcal{W} = \cup_n \mathcal{W}^n$ is an open cover of G that's plainly σ-locally finite and refines $\mathcal{U}$. Thus each open cover $\mathcal{U}$ of G admits a σ-locally finite open cover $\mathcal{W}$ that refines $\mathcal{U}$ so G is paracompact and therefore normal ([**?**], Theorem 5.28, Corollary 5.32).

Warning: *Not every topological group is normal as exhibited by the following classical example.*

EXAMPLE 3.0.11. *Let m be an uncountable cardinal number. Then $\mathbb{Z}^m$ is a T_0 group, hence completely regular. But $\mathbb{Z}^m$ is not normal.*

PROOF. We'll write $\mathbb{Z}^m$ as $\prod_{i \in I} \mathbb{Z}_i$ where each $\mathbb{Z}_i$ is $\mathbb{Z}$ and $|I| = m$. For the sake of the present efforts, let $A, B \subseteq \mathbb{Z}^m$ be given as follows

$$A = \{(x_i) \in \mathbb{Z}^m : \quad \text{for any} \quad n \neq 0, \quad \text{there is at most one index} \quad i \quad \text{for which} \quad x_i = n\},$$

and

$$B = \{(x_i) \in \mathbb{Z}^m : \quad \text{for any} \quad n \neq 1, \quad \text{there is at most one index} \quad i \quad \text{for which} \quad x_i = n\}.$$

Then A and B are disjoint. If $(x_i)_{i \in I} \notin A$ then there are $i_0, i_1 \in I, i_0 \neq i_1$ such that for some $n \in \mathbb{Z}, n \neq 0, x_{i_0} = x_{i_1} = n$. The set

$$\{(y_i)_{i \in I} \in \mathbb{Z}^m : y_{i_0} = y_{i_1} = n\}$$

is an open set containing $(x_i)_{i \in I}$ but no point of A, so A is a closed set. Similarly B is a closed set.

Let U, V be open subsets of $\mathbb{Z}^m$ such that $A \subseteq U$ and $B \subseteq V$. We claim that $U \cap V \neq \emptyset$, which will show that $\mathbb{Z}^m$ is not normal.

Let $(x_i^{(1)})_{i \in I} \in \mathbb{Z}^m$ be defined by $x_i^{(1)} = 0$ for each $i \in I$. Since $(x_i^{(1)}) \in A \subseteq U$, there are distinct indices $i_1, \ldots, i_{m_1} \in I$ such that

$$(x_i^{(1)})_{i \in I} \in \{(x_i)_{i \in I} \in \mathbb{Z}^m : x_{i_1} = x_{i_2} = \cdots x_{i_{m_1}} = 0\} \subseteq U.$$

Let $(x_i^{(2)})_{i\in I}$ be defined by

$$x_i^{(2)} = \begin{cases} k & \text{for} \quad i = i_k, k = 1, \ldots, m_1 \\ 0 & \text{otherwise.} \end{cases}$$

Since $(x_i^{(2)}) \in A \subseteq U$ there exists $i_{m_1+1}, \ldots, i_{m_2} \in I$, distinct indices from each other and from $i_1, \ldots, i_{m_1}$ such that

$$(x_i^{(2)})_{i\in I} \in \{(x_i)_{i\in I} \in \mathbb{Z}^m : x_{i_1} = 1, x_{i_2} = 2, \ldots, x_{i_{m_1}} = m_1, x_{i_{m_1}+1} = 0, \ldots, x_{i_{m_2}} = 0\} \subseteq U,$$

Continue in this manner.

Define $(y_i)_{i\in I} \in \mathbb{Z}^m$ as follows: $y_{i_k} = k$ for any k and $y_i = 1$ if $i \neq i_k$. Plainly, $(y_i)_{i\in I} \in B$. Hence for some finite subset J of I,

$$\{(x_i)_{i\in I} \in \mathbb{Z}^m : x_i = y_i, \quad \text{for} \quad i \in J\} \subseteq V.$$

But J is finite so there is an n_0 such that $i_k \notin J$ whenever $k > m_{n_0}$. Look at $(z_i)_{i\in I} \in \mathbb{Z}^m$, where

$$z_i = \begin{cases} k & \text{if} \quad i = i_k, k \leq m_{n_0} \\ 0 & \text{if} \quad i = i_k, m_{n_0} + 1 \leq k \leq m_{n_0+1} \\ 1 & \text{otherwise} \end{cases}$$

So

$$(z_i)_{i\in I} \in \{(x_i)_{i\in I} \in \mathbb{Z}^m : x_i = y_i, \quad \text{for} \quad i \in J\} \subseteq V.$$

So

$$(z_i)_{i\in I} \in \{(x_i)_{i\in I} \in \mathbb{Z}^m : x_{i_1} = 1, x_{i_2} = 2, \ldots, x_{i_{m_0}} = m_0, x_{i_{m_{n_0}}+1} = x_{i_{m_{n_0}}+2} = \cdots x_{i_{m_{n_0+1}}} = 0\} \subseteq U.$$

So $(z_i)_{i\in I} \in V, (z_i)_{i\in I} \in U$, and $U \cap V \neq \emptyset$. Therefore $\mathbb{Z}^m$ is not normal. □

After thoughts: The missing steps of the Birkhoff-Kakutani Theorem. To prove (3.1)

$$r < s \Rightarrow V_r \subseteq V_s,$$

suppose $s < 1$, and write r, s in dyadic forms

$$\begin{aligned} r &= 2^{-l_1} + 2^{-l_2} + \cdots + 2^{-l_n}, \ \ 0 < l_1 < l_2 < \cdots < l_n \\ s &= 2^{-m_1} + 2^{-m_2} + \cdots + 2^{-m_p}, \ \ 0 < m_1 < m_2 < \cdots < m_p, \end{aligned}$$

where $l_1, \ldots, l_m, m_1, \ldots, m_p$ are all positive integers. There must be a k so that

$$l_1 = m_1, l_2 = m_2, \ldots, l_{k-1} = m_{k-1} \ \text{ but } \ l_k > m_k.$$

Let

$$W = V_{2^{-l_1}} \cdot V_{2^{-l_2}} \cdot V_{2^{-l_{k-1}}}.$$

Then

$$\begin{aligned}
V_r &= W \cdot V_{2^{-l_k}} \cdot V_{2^{-l_k+1}} \cdots V_{2^{-l_n}} \\
&\subseteq W \cdot V_{2^{-l_k}} \cdot V_{2^{-l_k-1}} \cdot V_{2^{-l_k-2}} \cdots V_{2^{-l_n+1}} \cdot V_{2^{-l_n}} \cdot V_{2^{-l_n}} \\
&\subseteq W \cdot V_{2^{-l_k}} \cdot V_{2^{-l_k-1}} \cdot V_{2^{-l_k-2}} \cdots V_{2^{-l_n+1}} \cdot V_{2^{-l_n+1}} \cdot \\
&\quad \text{(since after all } V_{2^{-l_n}} \cdot V_{2^{-l_n}} = U_{l_n} \cdot U_{l_n} \subseteq U_{l_n-1} = V_{2^{-l_{n+1}}}) \\
&\vdots \\
&\subseteq W \cdot V_{2^{-l_k}} \cdot V_{2^{-l_k}} \subseteq W \cdot V_{2^{-l_k+1}} \\
&\subseteq W \cdot V_{2^{-m_k}} \subseteq V_{2^{-l_1}} \cdot V_{2^{-l_2}} \cdots V_{2^{-l_k-1}} \cdot V_{-m_k} \\
&\subseteq V_{2^{-m_1}} \cdot V_{2^{-m_2}} \cdots V_{2^{-m_{k-1}}} \cdot V_{2^{-m_k}} \\
&\subseteq V_{2^{-m_1}} \cdot V_{2^{-m_2}} \cdots V_{2^{-m_k}} \cdots V_{2^{-m_p}} \\
&= V_s.
\end{aligned}$$

With the same representation of r in mind we set off to prove (3.2)

$$V_r \cdot V_{2^{-l}} \subseteq V_{r+2^{-l+2}}.$$

We suppose that $r + 2^{-l+2} < 1$. If $l > l_n$ then

$$V_r \cdot V_{2-l} = V_{r+2^{-l}},$$

and all is well. So we look to the case that $l \leq l_n$.

Let k be the positive integer such that

$$l_{k-1} < l \leq l_k, \quad (l_0 = 0).$$

Let r_1 be given by

$$r_1 = 2^{-l+1} - 2^{-l_k} - 2^{-l_k+1} - \cdots - 2^{-l_n},$$

and

$$r_2 \leq r + r_1.$$

It's plain that

$$r < r_2 < r + 2^{-l+1},$$

so

$$V_r \cdot V_{2^{-l}} \subseteq V_{r_2} \cdot V^{2-l} = V_{r_2+2^{-l}} \subseteq V_{r+2^{-l+1}+2^{-l}} \subseteq V_{r+2^{-l+2}},$$

and that's that. □

CHAPTER 2

Lebesgue Measure in Euclidean Space

By an interval in $\mathbb{R}^n$ we mean any set I of the form

$$I = I_1 \times \cdots \times I_k$$

where $I_1, \cdots, I_k$ are finite intervals in $\mathbb{R}$. We do not ask that $I_1, \cdots, I_k$ all be open, closed or half-open/ half-closed mixtures are just fine. Each $I_j \subseteq \mathbb{R}$ has a length $l(I_j)$ and with this mind we define "volume" of I by

$$\mathrm{vol}(I) = \prod_{j \le k} l(I_j)$$

let $A \subseteq \mathbb{R}^k$. We define the **outer measure** or **Lebesgue outer measure** of A, $m^*(A)$ by

$$m^*(A) = \inf \left\{ \sum_n \mathrm{vol}(I_n) : I_n \text{ is an interval in } \mathbb{R}^k, A \subseteq \cup_n I_n \right\}.$$

Regarding edges: There is a great deal of latitude with regards to the nature of the edges of the intervals in the coverings of a set $A \subseteq \mathbb{R}^k$ that are used to compute $m^*(A)$.

For instance if we wish we can assume each edge has length less than δ for some fixed $\delta > 0$. This is plain since any interval I in $\mathbb{R}^k$ is the union of overlapping intervals all of whose edges have length less than δ and the sum of whose volume totals I's volume.

More, we can assume we're covering A by *open* intervals, that is, all the edges are open. In fact, if (I_j) is a covering of A by intervals and $\epsilon > 0$ then for each j we can enlarge I_j to an open interval J_j, $I_j \subseteq J_j$ and

$$\mathrm{vol}(J_j) < \mathrm{vol}(I_j) + \epsilon/2.$$

It follows that in computing $m^*(A)$, if (I_j) is a covering of A by intervals then we can find a sum $\sum \mathrm{vol}(J_j)$ that is as close as we please to $\sum \mathrm{vol}(I_j)$ where (J_j) is a covering of A by open intervals. Hence we can restrict our attention to finding the infimum of such sums $\sum \mathrm{vol}(J_j)$ where (J_j) is an open covering of A by intervals.

$\mathbf{1^0}$ *If* $A \subseteq B$ *then*

$$m^*(A) \le m^*(B).$$

$\mathbf{2^0}$ *If* $A = \cup_n A_n$ *then*

$$m^*(A) \le \sum_n m^*(A_n)$$

We can, and do, assume that $\sum_n m^*(A_n) < \infty$. Indeed, let $\epsilon > 0$ be given and choose for each n a sequence $(I_{n_j})$of intervals that cover A_n and satisfy

$$\sum_j \text{vol}(I_{n_j}) < m^*(A_n) + \frac{\epsilon}{2^{n_j}}.$$

Since $A = \cup_n A_n \subseteq \cup_{n,j} I_{n_j}$,

$$m^*(A) \le \sum_{n,j} \text{vol}(I_{n_j}) \le \sum_n \left(m^*(A_n) + \frac{\epsilon}{2^n}\right) \le \sum_n m^*(A_n) + \epsilon.$$

$\mathbf{3^0}$ *For any interval I,*

$$m^*(I) = vol(I).$$

Let $\epsilon > 0$ be given and let (I_j) be an open covering of I such that

$$\sum_j \text{vol}(I_j) < m^*(I) + \epsilon.$$

Take any closed subinterval J of I. Since J is compact there is a j_0 such that $J \subseteq I_1 \cup \cdots \cup I_{j_0}$. Let's look closely to the intervals $I_1, \cdots, I_{j_0}, J$. Each $(k-1)$ dimensional face of these intervals lies in a $(k-1)$ dimensional hyperplane in $\mathbb{R}^k$; in turn, these hyperplanes divide $\bar{I}_1, \ldots, \bar{I}_{j_0}$ into closed intervals $K_1, \ldots, K_{n_{j_0}}$; similarly J is divided into closed intervals $J_1, \ldots, J_{m_{j_0}}$ by the same hyperplane. (Think of the case $n = 3$.) Since $J \subseteq \cup_{j \le j_0} I_j$ each J_m is one of the K_n's so that

$$\begin{aligned}
\text{vol}(J) &= \sum_{m \le m_{j_0}} \text{vol}(J_m) \\
&\le \sum_{n \le n_{j_0}} \text{vol}(K_n) \\
&\le \sum_{j \le j_0} \text{vol}(I_j) \\
&\le m^*(I) + \epsilon.
\end{aligned}$$

This is so for every closed subinterval J of I so

$$\text{vol}(I) \le m^*(I) + \epsilon;$$

epsilonics soon tell us that

$$\text{vol}(I) \le m^*(I).$$

The reverse is plain.

Some ground work is needed to prepare the way for measurable sets.

A. *If F_1 and F_2 are disjoint closed bounded sets then*

$$m^*(F_1 \cup F_2) = m^*(F_1) + m^*(F_2).$$

Let $\delta > 0$ be chosen so that no interval of diameter less than δ meets both F_1 and F_2 (e.g., $\delta < \frac{1}{2}d(F_1, F_2)$.) Let $\epsilon > 0$ be given. Pick a sequence (I_i) of intervals of diameter less than δ such that

$$F_1 \cup F_2 \subseteq \cup_i I_i, \text{ and } \sum_i \text{vol}(I_i) \overset{\epsilon}{\sim} m^*(F_1 \cup F_2).$$

Denote by $(I_k^{(1)})$ those intervals among the I_i's that meet F_1 and by $(I_k^{(2)})$ those that meet F_2. $F_1 \subseteq \bigcup_j I_j^{(1)}$ and $F_2 \subseteq \bigcup I_j^{(2)}$ and by our judicious concerns over δ. Alas,

$$\begin{aligned} m^*(F_1) + m^*(F_2) &\leq \sum_j \text{vol}(I_j^{(1)}) + \sum_j \text{vol}(I_j^{(2)}) \\ &\leq \sum_i \text{vol}(I_i) \leq m^*(F_1 \cup F_2) + \epsilon; \end{aligned}$$

Epsilonics to the rescue: $m^*(F_1) + m^*(F_2) \leq m^*(F_1 \cup F_2)$.

B. *If G is a bounded open set then for each $\epsilon > 0$ there is a closed set $F \subseteq G$ such that*

$$m^*(F) > m^*(G) - \epsilon.$$

Represent $G = \cup_i I_i$ where I_i's are non-overlapping intervals, and let $\epsilon > 0$ be given; of course $m^*(G) \leq \sum_i \text{vol}(I_i)$ and so there is an n_0 so that

$$\sum_{i \leq n_0} \text{vol}(I_i) > m^*(G) - \frac{\epsilon}{2}.$$

For each $i \leq n_0$ let J_i be a closed subinterval of the interior of I_i with

$$\text{vol}(J_i) > \text{vol}(I_i) - \frac{\epsilon}{2^i}.$$

Then $F = \bigcup_{i \leq n_0} J_i$ is a closed subset of G and

$$\begin{aligned} m^*(F) &= m^*\left(\bigcup_{i \leq n_0} J_i\right) \\ &= \sum_{i \leq n_0} m^*(J_i) \;\; \text{(by A)} \\ &= \sum_{i \leq n_0} \text{vol}(J_i) \\ &> \sum_{i \leq n_0} \left(\text{vol}(I_i) - \frac{\epsilon}{2^i}\right) \\ &\geq \sum_{i \leq n_0} \text{vol}(I_i) - \frac{\epsilon}{2} \\ &> m^*(G) - \epsilon. \end{aligned}$$

NB: The openness of G was used to represent G in an appropriate way.

C. *If F is a closed subset of an open bounded set G then*

$$m^*(G \backslash F) = m^*(G) - m^*(F).$$

Let $\epsilon > 0$ be given. Using B, chose a closed set $F_1 \subseteq G\backslash F$ so that

$$m^*(F_1) > m^*(G\backslash F) - \epsilon.$$

Notice that

$$\begin{aligned} m^*(F) + m^*(G\backslash F) &\leq m^*(F) + (m^*(F_1) + \epsilon) \\ &= m^*(F_1 \cup F_2) + \epsilon \\ &\leq m^*(G) + \epsilon. \end{aligned}$$

Epsilonics take over to say

$$m^*(F) + m^*(G\backslash F) \leq m^*(G).$$

Of course, $\mathbf{2^0}$ takes care of the reverse inequality and with it, C.

A subset E of $\mathbb{R}^k$ is **Lebesgue measurable** if given an $\epsilon > 0$ there is a closed set F and open set G such that

$$F \subseteq A \subseteq G, \text{ and } m^*(G\backslash F) < \epsilon.$$

By the complementary nature of open and closed sets E *is Lebesgue measurable if and only if* E^c *is*:

$$F \subseteq E \subseteq G \Leftrightarrow G^c \subseteq E^c \subseteq F^c \text{ and } F^c\backslash G^c = G\backslash F.$$

$\mathbf{4^0}$ *If A and B are measurable then so is $A \cup B$.*

Pick F_A, F_B closed and G_A, G_B open such that

$$\begin{aligned} F_A &\subseteq A \subseteq G_A \text{ and } m^*(G_A\backslash F_A) \leq \tfrac{\epsilon}{2} \\ F_B &\subseteq B \subseteq G_B \text{ and } m^*(G_B\backslash F_B) \leq \tfrac{\epsilon}{2} \end{aligned}$$

$F = F_A \cap F_B$ is closed, $G = G_A \cap G_B$ is open, $F \subseteq A \cap B \subseteq G$, and $G\backslash F \subseteq (G_A\backslash F_A) \cup (G_B\backslash F_B)$ so that

$$m^*(G\backslash F) \leq m^*(G_A\backslash F_A) + m^*(G_B\backslash F_B) \leq \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.$$

$\mathbf{5^0}$ *A bounded set B is measurable if for each $\epsilon > 0$ there is a compact set $K \subseteq B$ such that*

$$m^*(K) > m^*(B) - \epsilon.$$

Suppose the bounded set B satisfies the conditions set forth and let $\epsilon > 0$ be given. We can find a compact set $K \subseteq B$ such that

$$m^*(K) \geq m^*(B) - \frac{\epsilon}{2}.$$

But $m^*(B) < \infty$ (why is that?). So we can cover B by a sequence (I_j) of open intervals each of diameter less than $\frac{1}{37}$ such that

$$\sum_j \text{vol}(I_j) < m^*(B).$$

Let G be the union of all those I_j's that meet B. $K \subseteq B \subseteq G$ and G is bounded. So our preparatory ... tells us that

$$\begin{aligned} m^*(G \backslash K) &= m^*(G) - m^*(K) \\ &\leq \sum_j m^*(I_j) - m^*(K) \\ &= \sum_j \text{vol}(I_j) - m^*(K) \\ &\leq m^*(B) + \frac{\epsilon}{2} - m^*(K) \\ &= m^*(B) - m^*(K) + \frac{\epsilon}{2} \\ &\leq \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon. \end{aligned}$$

$\mathbf{6^0}$ *(Finite) intervals are measurable.*

After all, if I is a (finite or bounded) interval $m^*(I) = \text{vol}(I)$ and so we can plainly approximate I from the inside by *compact* intervals.

$\mathbf{7^0}$ *Sets of outer measure zero are measurable.*

If $m^*(N) = 0$ and $\epsilon > 0$ is given there must be a sequence (I_j) of open intervals so $N \subseteq \cup_j I_j$ and

$$\sum_j \text{vol}(I_j) \leq m^*(N) + \epsilon = \epsilon.$$

Then $G = \cup_j I_j$ and $F = \emptyset$ soon show the way to N's measurability.

Another technical rest stop.

D. If (A_n) is a sequence of disjoint measurable sets of the interval I then $\cup_n A_n$ is measurable, too, and

$$m^*(\cup_n A_n) = \sum_n m^*(A_n).$$

Let $\epsilon > 0$ be given. Choose compact sets $F_n \subseteq A_n$ so

$$m^*(F_n) > m^*(A_n) - \frac{\epsilon}{2^{n+1}}.$$

Since

$$m^*(\cup_n A_n) \leq \sum_n m^*(A_n)$$

there is an $n \in \mathbb{N}$ so that

$$\sum_{n \leq n_0} m^*(A_n) > m^*(\cup_n A_n) - \frac{\epsilon}{2}.$$

If $F = \cup_{n \leq n_0}$ then F is compact (it's closed, and being a subset of I, bounded). Hence by A

$$m^*(F) = \sum_{n \leq n_0} m^*(F_n) > \sum_{n \leq n_0} m^*(A_n) - \frac{\epsilon}{2} > m^*(\cup_n A_n) - \frac{\epsilon}{2}.$$

We've just taken the bounded set $\cup_n A_n$ and for each $\epsilon > 0$ found a compact set K, contained in $\cup_n A_n$ so that

$$m^*(F) > m^*(A) - \epsilon.$$

$\cup_n A_n$ is measurable thanks to $\mathbf{5^0}$. Let's check the sums: for any $n \in \mathbb{N}$

$$\begin{aligned}\sum_{n\le m} m^*(A_n) &< \sum_{n\le m}\left(m^*(F_n) + \frac{\epsilon}{2^{n+1}}\right)\\ &\le \sum_{n\le m} m^*(F_n) + \frac{\epsilon}{2}\\ &= m^*\left(\cup_{n\le m} F_n\right) + \frac{\epsilon}{2}\\ &= m^*(\cup_n A_n) + \frac{\epsilon}{2}.\end{aligned}$$

Epsilonics assure us that

$$\sum_{n\le m} m^*(A_n) \le m^*(\cup_n A_n)$$

and this is so for each n. It follows that

$$\sum_n m^*(A_n) \le m^*\left(\bigcup_n A_n\right).$$

We've already seen the reverse so that's all she wrote for D.

$\mathbf{8^0}$ *If (A_n) is any sequence of disjoint measurable set of $\mathbb{R}^k$ then $\cup_n A_n$ is measurable and*

$$m^*(\cup_n A_n) = \sum_n m^*(A_n).$$

We'll bootstrap our way from D to $\mathbf{8^0}$. To start, let (I_m) be a sequence of disjoint intervals whose union is $\mathbb{R}^k$ and such that any bounded set in $\mathbb{R}^k$ is covered by finitely many I_m's.

For each $m, n \in \mathbb{N}$ let

$$A_{m,n} = I_m \cap A_n$$

be that part of A_n inside I_m. Each $A_{m,n}$ is measurable ($\mathbf{6^0}$ and $\mathbf{4^0}$) and then $A_{m,n}$'s are pairwise disjoint. Look to

$$\widetilde{A_m} = \cup_n A_{m,n},$$

the part of $\cup_n A_n$ in I_m. By D, $\widetilde{A_m}$ is measurable. Further, the $\widetilde{A_m}$'s are pairwise disjoint and $\cup_n \widetilde{A_m} = \cup_n A_n$.

Let $\epsilon > 0$ be given. For each m, choose a closed set $F_m \subseteq \widetilde{A_m}$ and an open set G_m, which is a bounded open set, $\widetilde{A_m} \subseteq G_m$ such that

$$m^*(G_m \backslash F_m) < \frac{\epsilon}{2^m}.$$

Look at $F = \cup_m F_m$ and $G = \cup_m G_m$. F is closed (if (x_n) is a convergent sequence of points of F then (x_n) is bounded and so for some m_0, (x_n) is a sequence on $\cup_{m\le m_0} F_m$ hence visits one of

$F_1, \ldots, F_{m_0}$ infinitely often - which ever F_j it visits so often contains its limit and G is open.

$$F = \cup_m F_m \subseteq \cup_m \widetilde{A_m} = \cup_m A_m = \cup_m \widetilde{A_m} \subseteq \cup_m G_m = G.$$

Further,

$$G \backslash F = \cup_m (G_m \backslash F) \subseteq \cup_m (G_m \backslash F_m)$$

so

$$m^*(G \backslash F) \leq \sum_m m^*(G_m \backslash F_m) \leq \sum_m \frac{\epsilon}{2^m} = \epsilon,$$

and $\cup_n A_n$ is measurable.

Now

$$A_n = \cup A_{m,n}$$

so

$$m^*(A_n) \leq \sum_m m^*(A_{m,n}).$$

It follows that

$$\begin{aligned} \sum_n m^*(A_n) &\leq \sum_n \sum_m m^*(A_{m,n}) \\ &= \sum_m \sum_n m^*(A_{m,n}) \\ &= \sum_m m^*(\widetilde{A_m}), \end{aligned}$$

by D. Take $m \in \mathbb{N}$. Then

$$\begin{aligned} \sum_{j \leq m} m^*(\widetilde{A_j}) &\leq \sum_{j \leq m} (m^*(F_j) + m^*(G_j \backslash F_j)) \\ &= m^* \left(\bigcup_{j \leq m} F_j \right) + \sum_{j \leq m} m^*(G_j \backslash F_j) \quad \text{(by A)} \\ &\leq m^* \left(\bigcup_{j \leq m} F_j \right) + \sum_{j \leq m} \frac{\epsilon}{2^j} \\ &\leq m^* \left(\bigcup_n A_n \right) + \epsilon. \end{aligned}$$

The usual epsilonics leads us to conclude that

$$\sum_m m^*(\widetilde{A_m}) \leq m^* \left(\bigcup_n A_n \right)$$

and in tandem with what gone on before we see

$$\sum_n m^*(A_n) \leq m^* \left(\bigcup_n A_n \right).$$

Again the reverse holds without assumption so $\mathbf{8^0}$ is proved.

THEOREM 0.0.12 (The Fundamental Theorem of Lebesgue Measure). *In summary*

(i) m^* *is a non-negative, extended real-valued function defined for every subset of* $\mathbb{R}^k$ *which assigns*
- *to each interval, a value equal to its volume,*
- *to each set, a value common to all its translates,*
- *to bigger sets, bigger values*
- *to compact sets, finite values*
- *to non-empty sets, positive values*

and is countably subadditive in doing so. For any $A \subseteq \mathbb{R}^k$,

$$m^*(A) = \inf\{m^*(G) : G \text{ is open } A \subseteq G\}.$$

(ii) *The Lebesgue measurable subsets of* $\mathbb{R}^k$ *form a* $\sigma-$*field* $\mathcal{M}$ *of sets containing every open set, closed set, interval, and set of outer measure zero;* $E \in \mathcal{M}$ *if and only if* E*'s translates are members of* $\mathcal{M}$.

(iii) m^* *is countably additive on* $\mathcal{M}$ *and for* $E \in \mathcal{M}$,

$$m^*(E) = \sup\{m^*(K) : K \text{ is compact }, K \subseteq E\}.$$

CHAPTER 3

Invariant Measures on $\mathbb{R}^n$

1. Introduction

As we proceed in our study of invariant measures we will encounter theorems that assert the uniqueness of such measures in varying degrees of generality and in differing senses.

On compact object the uniqueness will be with regard to **Borel probabilities** (positive Borel measures of total mass 1).

If G is a locally compact group then we'll see that there is a unique left invariant regular Borel measure acting on (the Borel subsets of G) with uniqueness taken to mean 'up to multiplicative constants.'

In concrete cases it occurs that those multiplicative constants can themselves be of considerable interest, representing the rate of exchange between various natural geometric view points of the groups under consideration.

In this chapter we encounter an early example of seemingly mulitple invariant measures on a concrete group and compute the constants that allow one to convert from one viewpoint to another.

Our setting with be $\mathbb{R}^n$. We will be dealing with Lebesgue $n-$measure λ_n on $\mathbb{R}^n$ and Hausdorff $n-$measure $\mu^{(n)}$ on $\mathbb{R}^n$. Each is translation invariant and, sure enough, for each n, there is a constant κ_n such that

$$\mu^{(n)} = \kappa_n \lambda_n.$$

The constant κ_n is in fact the precise rate of exchange that allows one to move from a rectalinear view of $\mathbb{R}^n$ (as expressed by λ_n) to a spherical view (as found in $\mu^{(n)}$.)

Along the way to establishing the existence of κ_n and of computing it, we meet some of the real stalwarts of measure theory.

In the first section we give the elegant Hadwiger-Ohlmann proof of the **Brunn-Minkowski Theorem**, a geometric version of the Arthimetic-Geometric Mean Inequality. As a simple consequence of this we derive the **Isodiametric Inequality** which says that among Borel sets in $\mathbb{R}^n$ of the same diameter, the ball has the greatest volume.

In the next section we derive the still-wonderful **Covering Theorem of Vitali** for Vitali families of balls. We follow this with a short introduction to Hausdorff measure on $\mathbb{R}^n$, $\mu^{(n)}$. This leads to the main course of this chapter, the proof that for each n there is a $\kappa_n > 0$ such that

$$\mu^{(n)} = \kappa_n \lambda_n.$$

2. The Brunn-Minkowski Theorem

THEOREM 2.0.13 (The Brunn-Minkowski Theorem). *Let $n \geq 1$, and let λ_n denote Lebesgue measure on $\mathbb{R}^n$. If A and B are compact subsets of $\mathbb{R}^n$ then*

$$(\lambda_n(A+B))^{1/n} \geq (\lambda_n(A))^{1/n} + (\lambda_n(B))^{1/n} \quad (BM)$$

where

$$A + B = \{a + b : a \in A, b \in B\}.$$

Notice that (BM) is a geometric generalization of the Arithmetic-Geometric Mean Inequality for if A and B are rectangles with sides of length $(a_j)_{j=1}^n$ and $(b_j)_{j=1}^n$ respectively, then (BM) looks like

$$\left[\prod_i^n (a_j + b_j)\right]^{1/n} \geq \left(\prod_1^n a_j\right)^{1/n} + \left(\prod_1^n b_j\right)^{1/n}. \quad (\mathrm{BM})'.$$

Homogeneity lets us reduce this to the case where $a_j + b_j = 1$ for each j. But now the Arithmetic-Geometric Mean Inequality assures that

$$\frac{1}{n}\sum_{j=1}^n a_j \geq \left(\prod_1^n a_j\right)^{1/n}, \text{ and } \frac{1}{n}\sum_{j=1}^n b_j \geq \left(\prod_1^n b_j\right)^{1/n}.$$

So (noting that $\sum_{j=1}^n a_j + \sum_{j=1}^n b_j = \sum_{j=1}^n (a_j + b_j) = \sum_{j=1}^n 1 = n$)

$$1 = \frac{1}{n} \cdot n \geq \left(\prod_1^n a_j\right)^{1/n} + \left(\prod_1^n b_j\right)^{1/n},$$

which is $(\mathrm{BM})'$. Thus we have proved (BM) for boxes, rectangular parallelpipeds whose sides are parallel to the coordinate hyperplanes.

To prove (BM), note that A and B each are the union of finitely many rectangles whose interiors are distinct. We proceed by induction on the total number of rectangles in A and B. It is important to realize that the inequality is inaffected if we translate A and B independently: in fact, replacing A by $A + h$ and B by $B + k$ replaces $A + B$ by $A + B + h + k$ and the corresponding measures are the same as what we started with. If R_1 and R_2 are essentially disjoint rectangles in the collection making up A then they can be separated (a translation may be necessary) by a coordinate hyperplane $\{x_j = 0\}$ say. Thus we may assume that R_1 lies in $A^- = A \cap \{x_j \leq 0\}$ and R_2 lies in $A^+ = A \cap \{x_j \geq 0\}$. Notice that each A^+ and A^- contain at least one rectangle less than does A and $A = A^+ \cup A^-$.

What to do with B? Well, slide B over so if $B^+ = B \cap \{x_j \geq 0\}$ then

$$\frac{\lambda_n(B^+)}{\lambda_n(B)} = \frac{\lambda_n(A^+)}{\lambda_n(A)};$$

of course this entails

$$\frac{\lambda_n(B^-)}{\lambda_n(B)} = \frac{\lambda_n(A^-)}{\lambda_n(A)}$$

as well. But

$$(A^+ + B^+) \cup (A^- + B^-) \subseteq A + B$$

and the union on the left hand side is essentially disjoint. Moreover the total number of rectangles in either A^+ and B^+ or in A^- and B^- is less than that in A and B. Our induction hypothesis applies.

The result?

$$\begin{aligned}
\lambda_n(A+B) &\geq \lambda_n(A^+ + B^+) + \lambda_n(A^- + B^-) \\
&\geq (\lambda_n(A^+)^{1/n} + \lambda_n(B^+)^{1/n})^n + ((\lambda_n(A^-)^{1/n} + \lambda_n(B^-)^{1/n})^n \quad \text{(induction hypothesis)} \\
&= \lambda_n(A^+)\left(1 + \left(\frac{\lambda_n(B)}{\lambda_n(A)}\right)^{1/n}\right)^n + \lambda_n(A^-)\left(1 + \left(\frac{\lambda_n(B)}{\lambda_n(A)}\right)^{1/n}\right)^n \\
&= \lambda_n(A)\left(1 + \left(\frac{\lambda_n(B)}{\lambda_n(A)}\right)^{1/n}\right)^n \\
&= (\lambda_n(A)^{1/n} + \lambda_n(B)^{1/n})^n
\end{aligned}$$

and that's all she wrote. Thus we have (BM) for finite unions of boxes.

If A and B are open sets of finite measure then once given a margin of error, $\epsilon > 0$, we can find unions A_ϵ, B_ϵ of essentially disjoint rectangles such that $A_\epsilon \subseteq A, B_\epsilon \subseteq B$, and

$$\lambda_n(A) \leq \lambda_n(A_\epsilon) + \epsilon,\ \lambda_n(B) \leq \lambda_n(B_\epsilon) + \epsilon.$$

Once done, $A_\epsilon + B_\epsilon \subseteq A + B$ and we can apply the work of the previous paragraphs:

$$\begin{aligned}
\lambda_n(A+B)^{1/n} &\geq \lambda_n(A_\epsilon + B_\epsilon)^{1/n} \\
&\geq (\lambda_n(A) - \epsilon)^{1/n} + (\lambda_n(B) - \epsilon)^{1/n}.
\end{aligned}$$

Since $\epsilon > 0$ was but a margin of error and we can assume arbitrary small errors and made at worst, it follows that (BM) holds for open sets A and B of finite measure.

If A and B are compact sets then $A + B$ is compact as well. Look at $A_\epsilon = [d(x, A) < \epsilon]$. Then A_ϵ is open, contains A and $A_\epsilon \searrow A$. Similarly B_ϵ is defined analogously. What can we say about $(A + B)_\epsilon$? The same conclusions can be drawn. Further

$$A + B \subseteq A_\epsilon + B_\epsilon \subseteq (A + B)_{2\epsilon}.$$

Applying (BM) to A_ϵ, B_ϵ we get

$$\begin{aligned}
\lambda_n((A+B)_{2\epsilon})^{1/n} &\geq \lambda_n(A_\epsilon + B_\epsilon)^{1/n} \\
&\geq \lambda_n(A_\epsilon)^{1/n} + \lambda_n(B_\epsilon)^{1/n} \\
&\geq \lambda_n(A)^{1/n} + \lambda_n(B)^{1/n}.
\end{aligned}$$

Letting $\epsilon \to 0$ gives (BM) for $A, B, A + B$.

The general situation of A, B and $A + B$ measurable follows by approximating from within: if $\tilde{A}, \tilde{B}$ are compact sets with $\tilde{A} \subseteq A, \tilde{B} \subseteq B$ then $\tilde{A} + \tilde{B}$ is a compact set inside $A + B$, and

$$\lambda_n(A+B) \geq \lambda_n(\tilde{A} + \tilde{B}) \geq (\lambda_n(\tilde{A})^{1/n} + \lambda_n(\tilde{B})^{1/n})^n.$$

Let $\lambda_n(\tilde{A}) \nearrow \lambda_n(A), \lambda_n(\tilde{B}) \nearrow \lambda_n(B)$ and be done with it. $\square$

With the Brunn-Minkowski inequality in hand we can easily find our way to the **isodiametric inequality**.

THEOREM 2.0.14. *For any Borel set $B \subseteq \mathbb{R}^n$ we have*

$$\lambda_n(B) \leq \left(\frac{diam\ B}{2}\right)^n \lambda_n(B^n),$$

where as usual B^n denotes the closed unit ball of $\mathbb{R}^n$.

PROOF. Without loss of sleep we can assume that

$$d := \text{diam } B < \infty.$$

This in mind, realize that if $x, y, \in B$ then $||x - y|| \leq d$. Hence

$$B - B \subseteq d\, B^n,$$

and so

$$\begin{aligned} 2\lambda_n(B)^{1/n} &= \lambda_n(B)^{1/n} + \lambda_n(B)^{1/n} \\ &= \lambda_n(B)^{1/n} + \lambda_n(-B)^{1/n} \\ &\leq \lambda_n(B - B)^{1/n} \text{ (Brunn-Minkowski to the rescue!)} \\ &\leq \lambda_n(dB^n)^{1/n} \\ &= d\lambda_n(B^n)^{1/n}. \end{aligned}$$

It's easy to deduce the conclusion of the theorem from this inequality:

$$2\lambda_n(B)^{1/n} \leq \text{diam}(B)\lambda_n(B^n)^{1/n}. \quad \square$$

3. Vitali's Covering Theorem

THEOREM 3.0.15 (Vitali). *Let $\mathcal{F}$ be a family of closed balls in $\mathbb{R}^n$ that covers a set E in the sense of Vitali. i.e., given an $\epsilon > 0$ and $x \in E$ there is a $B \in \mathcal{F}$ such that the diameter of B is less than ϵ and $x \in B$. Then $\mathcal{F}$ contains a disjoint sequence that covers λ_n-almost all of E.*

PROOF. First we suppose E is bounded and contained in the bounded open set G. We disregard any members of $\mathcal{F}$ that aren't contained in G as well as those that don't intersect E. The resulting family, which we will still refer to as $\mathcal{F}$ covers E in the sense of Vitali. Our proof will be by 'exhaustion.' To start let R_1 be

$$R_1 := \sup\{\text{radius}(B) : B \in \mathcal{F}\}.$$

Choose $B_1 \in \mathcal{F}$, say centered at a, with radius r_1 so that

$$\frac{R_1}{2} \leq r_1.$$

Next let R_2 be

$$R_2 = \sup\{\text{radius } (B) : B \in \mathcal{F}, B_1 \cap B_2 = \emptyset\}.$$

Choose $B_2 \in \mathcal{F}$, say centered at a_2 with radius r_2 so that $B_2 \cap B_1 = \emptyset$ and

$$\frac{R_2}{2} \leq r_2.$$

Continue down this primrose lane. After k steps, let R_{k+1} be

$$R_{k+1} = \sup\{\text{radius}(B) : B \in \mathcal{F}, B \cap (B_1 \cup \cdots \cup B_k) = \emptyset\}.$$

Choose $B_{k+1} \in \mathcal{F}$ with center a_{k+1}, radius r_{k+1} so

$$B_{k+1} \cap (B_1 \cup \cdots \cup B_k) = \emptyset$$

and

$$\frac{R_{k+1}}{2} \leq r_{k+1}.$$

Naturally,

$$R_1 \geq R_2 \geq \cdots$$

More is so. The B_k's are pairwise disjoint so

$$\sum_n \lambda_n(B_k) = \lambda_n\left(\bigcup_k B_k\right) \leq \lambda_n(G) < \infty;$$

hence

$$\lim_k \lambda_n(B_k) = 0.$$

But

$$\lambda_n(B_k) = c_n r_k^n$$

(and we know c_n but we don't need to right now) so

$$\lim_k r_k = 0.$$

Since $0 \leq R_k \leq 2r_k$ we see

$$\lim_k R_k = 0.$$

The B_k's eat up E. How? What does it mean for $x \in E$ *not* to be in $\bigcup_{k \leq K} B_k$? Well $x \notin \bigcup_{k \leq K} B_k$ means

$$d\left(x, \bigcup_{k \leq K} B_k\right) = \delta > 0.$$

Choose $B \in \mathcal{F}$ so $x \in B$ and B (which is centered at say a with radius r) has radius less than $\delta/2$.

Each point of B is within $2r$ of x and $2r < \delta$. Hence B is disjoint from $B_1 \cup \cdots \cup B_K$. B is one of those balls belonging to $\mathcal{F}$ that take part in defining R_{k+1}; in particular, $r \leq R_{k+1}$. But

$$\lim_j R_j = 0$$

and the R_j's are monotone and non-increasing. Hence there is a j (which is necessarily bigger than K) so that

$$R_{j+1} < r \leq R_j.$$

Since $r > R_{j+1}$, B must meet one of the balls $B_1, \ldots, B_j$; Suppose $B \cap B_m \neq \emptyset$. Since B does *not* meet $B_1 \cup \cdots \cup B_K$, $m > K$. It follows that for any $c \in B \cap B_m$

$$\begin{aligned} ||x - a_m|| &\leq ||x - a|| + ||a - c|| + ||c - a_m|| \text{ (where } a_m \text{ is the center of } B_m \text{ and } a \text{ is the center of } B) \\ &\leq r + r + r_m \\ &= 2r + r_m. \end{aligned}$$

So

$$||x - a_m|| \leq 2r + r_m,$$

where

$$r \leq R_j \leq R_m \leq 2r_m.$$

WHY?

Hence

$$||x - a_m|| < 2r + r_m \leq 2 \cdot (2r_m) + r_m \leq 5r_m.$$

To summarize: if $x \in E \setminus \cup_{k \leq K} B_k$ then x belongs to the union $\cup_{k>K} \tilde{B}_k$ of those closed balls $\tilde{B}_k$, where $\tilde{B}_k$ shares the center a_k with B_k but $\tilde{B}_k$ has five times the radius of B_k.

In other words,

$$\begin{aligned} \lambda_n\left(E \setminus \bigcup_{k \leq K} B_k\right) &\leq \sum_{k \leq K} \lambda_n(\tilde{B}_k) \\ &= \sum_{k>K} 5^n \lambda_n(B_k) \to 0, \text{ as } K \to \infty, \end{aligned}$$

since, as we have already noted,

$$\sum_k \lambda_n(B_k) = \lambda_n\left(\bigcup_k B_k\right) \leq \lambda_n(G) < \infty.$$

For general E, look at the sets

$$E_m = \{x \in E : m - 1 \leq ||x|| < m\}, \ m \in \mathbb{N}.$$

Of course,

$$\lambda_n(E \setminus \bigcup_m E_m) = 0$$

and each E_m is as we've discussed above. What's more, the G we chose at the very start of our proof can be chosen to be the open set

$$\{x \in \mathbb{R}^n : m - 1 < ||x|| < m\}$$

so the disjointness is achieved by passing from one E_m to the next. □

4. Hausdorff $n-$measure on $\mathbb{R}^n$

A Hausdorff gauge function is a map

$$h : [0, \infty] \to [0, \infty]$$

such that $h(0) = 0, h(t) > 0$ whenever $t > 0$, $h(\infty) \leq \infty$, h is ascending and right-continuous from $[0, \infty)$ to $[0, \infty)$.

Let (X, ρ) be a metric space. For $E \subseteq X$ let $h(E) = h(\text{diam } E)$, $h(\emptyset) = 0$. Notice that h so defined is a premeasure on any class $\mathcal{C}$ containing $\emptyset$. Let $\delta > 0$. Consider the following classes $\mathcal{C}$ of subsets of X:

$$\begin{aligned} \mathcal{C} &= \mathcal{G} = \{G \subseteq X : G \text{ is open}\} \\ \mathcal{C} &= \mathcal{F} = \{F \subseteq X : F \text{ is closed}\} \\ \mathcal{C} &= \mathcal{P} = \{S \subseteq X\}. \end{aligned}$$

On each of these classes, we have intermediate measures

$$\mu^h_\delta, \nu^h_\delta, \sigma^h_\delta, \text{ and } \tau^h_\delta,$$

$$\begin{aligned} \mu^h_\delta(E) &:= \inf\left\{\sum_n h(G_n) : G_n \in \mathcal{G}, E \subseteq \cup_n G_n, \text{diam } G_n \le \delta\right\}, \\ \nu^h_\delta(E) &:= \inf\left\{\sum_n h(F_n) : F_n \in \mathcal{F}, E \subseteq \cup_n F_n, \text{diam } F_n \le \delta\right\}, \\ \sigma^h_\delta(E) &:= \inf\left\{\sum_n h(S_n) : E \subseteq \cup_n S_n, \text{diam } S_n \le \delta\right\}, \\ &\text{and} \\ \tau^h_\delta(E) &:= \inf\left\{\sum_n h(S_n) : E = \cup_n S_n, \text{diam } S_n \le \delta\right\}. \end{aligned}$$

PROPOSITION 4.0.16. *If $0 < \delta < \epsilon$ then*

$$\mu^h_\epsilon \le \nu^h_\delta = \sigma^h_\delta = \tau^h_\delta \le \mu^h_\delta.$$

COROLLARY 4.0.17. *If μ^h, ν^h and σ^h are the measures derived from h and the classes $\mathcal{G}, \mathcal{F}$ and $\mathcal{P}$ a la Method II then*

$$\mu^h = \nu^h = \sigma^h.$$

Moreover

$$\mu^R(E) = \sup_{\delta>0} \tau^h_\delta(E).$$

As C.A. Rogers noted, this shows that the μ^h measure of E can be defined in terms of the diameters and covering properties of *subsets* of E. Thus the μ^h-measure of E is an intrinsic property of E as a set with the metric ρ. So $\mu^h(E)$ doesn't change when E is put (isometrically) insider another metric space (X', ρ') than (X, ρ).

PROOF. Several assertions of the proposition are clear or, at the very least, easy-to-see. So ν^h_δ requires covering by *closed* sets of diameter $\le \delta$ and so $\sigma^h_\delta \le \nu^h_\delta$. On the other hand, diam $S =$ diam $\bar{S}$ so in fact, $\sigma^h_\delta = \nu^h_\delta$.

Again it's plain that $\sigma^h_\delta \le \tau^h_\delta$; yet if (S_n) is a sequence of sets that covers E with each S_n having diameter $\le \delta$ and $T_n = S_n \cap E$ then $\{T_n\}$ also covers E, diam $T_n \le \delta$ and $\cup_n T_n = E$; it soon follows that

$$\tau^h_\delta(E) \le \sum_n h(T_n) \le \sum_n h(S_n),$$

and so $\tau^h_\delta \leq \sigma^h_\delta$. In sum:

$$\tau^h_\delta = \sigma^h_\delta.$$

Naturally, $\sigma^h_\delta \leq \mu^h_\delta$ and so the only missing ingredient to a complete proof is that $\mu^h_\epsilon \leq \nu^h_\delta$ when $\delta < \epsilon$. Here some epsilonics are called for. Let $\eta > 0$. Cover E by a sequence (F_n) of closed sets with diameter $\leq \delta$ so that

$$\nu^h_\delta(E) \leq \sum_n h(F_n) \leq \nu^h_\delta(E) + \eta$$

(we clearly need only consider what happens when $\nu^h_\delta(E) < \infty$).

Choose $\eta > 0$ so that

$$h(\text{diam } F_n + 2\eta_n) < h(F_n) + \frac{\eta}{2^n}$$

and

$$\delta + 2\eta_n < \epsilon.$$

Let

$$U_n = [d(x, F_n) < \eta_n];$$

each U_n is open, contains F_n and

$$\text{diam } U_n \leq \text{diam } F_n + \eta_n + \eta_n \leq \delta + 2\eta_n < \epsilon.$$

Further

$$h(U_n) \leq h(\text{diam } + 2\eta_n) \leq h(F_n) + \frac{\eta}{2^n}.$$

Hence $E \subseteq \cup_n U_n$ each U_n open, has diameter less than ϵ and

$$\begin{aligned}\sum_n h(U_n) &\leq h(\text{diam } F_n + 2\eta_n) \\ &\leq \sum_n h(F_n) + \frac{\eta}{2^n} \leq \nu^h_\delta(E) + \eta.\end{aligned}$$

It follows that

$$\mu^h_\epsilon(E) \leq \nu^h_\delta + \eta.$$

Since $\eta > 0$ was arbitrary, the proof is complete. □

Let $h(t) = t^n$. The Hausdorff measure generated by h is denoted by $\mu^{(n)}$. Keep in mind that 'the Hausdorff measure' is a misnomer; after all $\mu^{(n)}$ makes sense in any metric space (X, ρ).

A comparison of $\mu^{(n)}$ and $n-$dimensional Lebesgue measure λ_n on $\mathbb{R}^n$ is worth our while. The comparison proceeds in two steps. In the first, we'll argue to their proportionality, in the second step we'll complete the constant of proportionality.

THEOREM 4.0.18. *Let*

$$C_0 = \{x = (x_1, \ldots, x_n) \subseteq \mathbb{R}^n : 0 \leq x_i < 1, i = 1, 2, \ldots, n\}$$

and

$$\kappa_n = \mu^{(n)}(C_0).$$

Then

$$\mu^{(n)}(E) = \kappa_n \lambda_n(E).$$

PROOF. Both $\mu^{(n)}$ and λ_n are invariant under translations and are homogeneous of order n. (i.e., for any $t \geq 0$, $\mu^{(n)}(tE) = t^n \mu^{(n)}(E)$ and $\lambda_n(tE) = t^n \lambda_n(E)$.) It follows that

$$\mu^{(n)}(C) = \kappa_n \lambda_n(C)$$

for any cube of the form

$$C = \{x = (x_1, \ldots, x_n) : a_i \leq x_i < a_i + s\}$$

where $(a_1, \ldots, a_n) \in \mathbb{R}$ and $s > 0$.

Consider the collection $\mathcal{C}$ of cubes of the form

$$\left\{x = (x_1, \ldots, x_n) : \frac{r_{i-1}}{2^n} \leq x_i < \frac{r_i}{2^n}\right\}$$

where $k \geq 0$ is an integer and $r_1, \ldots, r_n \in \mathbb{Z}$. $\mathcal{C}$ enjoys the following property: if two members of $\mathcal{C}$ have a point in common, then one is a subset of the other.

So what? Well if G is any open set then each point $g \in G$ lies in a cube $C \in \mathcal{C}$ that is contained entirely within G and is maximized in the sense that C is not a subset of any larger member of $\mathcal{C}$ that is entirely inside G. Hence G is the union of the maximal cubes from $\mathcal{C}$ and these cubes are disjoint. Let's write

$$G = \cup_m C_m,$$

where (C_m) is the sequence of maximal cubes just described. Everything we're talking about is measurable and so

$$\mu^{(n)}(G) = \sum_m \mu^{(n)}(C_m) = \sum_m \kappa_n \lambda_n(C_m) = \kappa_n \lambda_n(G).$$

So

$$\mu^{(n)}(G) = \kappa_n \lambda_n(G)$$

for open sets $G \subseteq \mathbb{R}^n$.

Next we bootstrap this equality's verity to $\mathcal{G}_\delta-$sets. Let H be a $\mathcal{G}_\delta-$ subset of $\mathbb{R}^n$. If $\mu^{(n)}(H)$ and $\lambda^{(n)}(H)$ are both equal to $+\infty$ then there's little that need be said. Naturally, they're also finite at the same time. In fact, if $\mu^{(n)}(H) < \infty$ then we can cover H by a sequence (G_m) of open sets with $\sum_n (\text{diam } G_m)^n < \infty$; if we enclose each G_m in an open cube U_m where each edge of U_m is twice the diameter of G_m then $H \subseteq \cup_m U_m$ and

$$\begin{aligned} \lambda_n(\cup_m C_m) &\leq \sum_m \lambda_n(C_m) \leq \sum_m (2 \cdot \text{diam } G_m)^n \\ &\leq 2^n \sum_m (\text{diam } G_m)^n < \infty. \end{aligned}$$

On the other hand, if $\lambda_n(H) < \infty$ then H can be covered by an open set of finite λ_n measure and so (since for open sets $\mu^{(n)}$ and $\kappa_n \lambda_n$ agree),

$$\mu^{(n)}(H) \leq \mu^{(n)}(G) = \kappa_n \lambda_n(G)$$

for any appropriate open set $H \subseteq G$ and $\mu^{(n)}(H) < \infty$ follows.

Okay, suppose $\mu^{(n)}(H)$ and $\kappa_n\lambda_n(H)$ are both finite. H is a $\mathcal{G}_\delta$ set so we can write H in the form

$$H = \cap_m G_m,$$

where (G_m) is a descending sequence of open sets each of finite measure (be that measure $\mu^{(n)}$ or λ_n!). It follows that

$$\mu^{(n)}(H) = \inf_m \mu^{(n)}(G_m) = \inf_m \kappa_n\lambda_n(G_m) = \kappa_n\lambda_n(H),$$

thanks to the measurability of all sets involved. So

$$\mu^{(n)}(H) = \kappa_n\lambda_n(H)$$

for any $\mathcal{G}_\delta$−set $H \subseteq \mathbb{R}^n$.

Now look at any $E \subseteq \mathbb{R}^n$. Choose $\mathcal{G}_\delta$− sets $H_1, H_2 \in \mathbb{R}^n$ with $E \subseteq H_1 \cap H_2$ so

$$\mu^{(n)}(E) = \mu^{(n)}(H_1), \text{ and } \lambda_n(E) = \lambda_n(H_2).$$

Of course

$$\mu^{(n)}(E) \leq \mu^{(n)}(H_1 \cap H_2) \leq \mu^{(n)}(H_1) = \mu^{(n)}(E)$$

and

$$\lambda_n(E) \leq \lambda_n(H_1 \cap H_2) \leq \lambda_n(H_2) = \lambda_n(E)$$

so

$$\mu^{(n)}(E) = \mu^{(n)}(H_1 \cap H_2) \text{ and } \lambda_n(E) = \lambda_n(H_1 \cap H_2).$$

It soon follows that

$$\mu^{(n)}(E) = \mu^{(n)}(H_1 \cap H_2) = \kappa_n\lambda_n(H_1 \cap H_2) = \kappa_n\lambda_n(E)$$

since $H_1 \cap H_2$ is a $\mathcal{G}_\delta$ set. □

THEOREM 4.0.19.

$$\kappa_n = \left(\frac{4}{\pi}\right)^{n/2} \Gamma\left(\frac{n+2}{2}\right).$$

PROOF. To compute κ_n we need to know the volume, $\lambda_n(\overline{B})$, of the closed unit ball in $\mathbb{R}^n$; cutting to the quick, this is given by

$$\lambda_n(\overline{B}) = \frac{\pi^{n/2}}{\Gamma\left(\frac{n+2}{2}\right)}.$$

Let's see what is involved. Let (S_k) be a sequence of subsets of $\mathbb{R}^n$ that cover the cube C_0. Then

$$\begin{aligned} 1 = \lambda_n(C_0) \leq \lambda_n(\cup_k S_k) & \\ &\leq \sum_k \lambda_n(S_k) \\ &\leq \sum_k \left(\frac{\pi}{4}\right)^{n/2} \frac{(\operatorname{diam}(S_k))^n}{\Gamma\left(\frac{n+2}{2}\right)} \end{aligned}$$

by the isodiametric inequality. It follows that

$$1 \leq \left(\frac{\pi}{4}\right)^{n/2} \frac{\mu^{(n)}(C_0)}{\Gamma\left(\frac{n+2}{2}\right)} = \left(\frac{\pi}{4}\right)^{n/2} \frac{\kappa_n}{\Gamma\left(\frac{n+2}{2}\right)}.$$

On the other hand, if $\delta > 0$ is given to us then Vitali's covering theorem provides us with a sequence (B_k) of disjoint balls each of diameter less than δ and each contained in C_0 with

$$\lambda_n(C_0) = \sum_k \lambda_n(B_k), \quad \text{and} \quad \lambda_n(C_0 \setminus \cup_k B_k) = 0.$$

It follows that

$$\mu^{(n)}(C_0 \setminus \cup_k B_k) = 0$$

Hence

$$\begin{aligned}
\mu_\delta^{(n)}(C_0) &\leq \mu_\delta^{(n)}(\cup_k B_k) + \mu_\delta^{(n)}(C_0 \setminus \cup_k B_k) \\
&= \mu_\delta^{(n)}(\cup_k B_k) \\
&\leq \sum_k (\text{diam } B_k)^n \\
&= \sum_k \left(\frac{4}{\pi}\right)^{n/2} \Gamma\left(\frac{n+2}{2}\right) \lambda_n(B_k) \\
&= \left(\frac{4}{\pi}\right)^{n/2} \Gamma\left(\frac{n+2}{2}\right) \sum_k \lambda_n(B_k) \\
&= \left(\frac{4}{\pi}\right)^{n/2} \Gamma\left(\frac{n+2}{2}\right) \lambda_n(C_0) \\
&= \left(\frac{4}{\pi}\right)^{n/2} \Gamma\left(\frac{n+2}{2}\right).
\end{aligned}$$

So

$$\kappa_n = \mu^{(n)}(C_0) \leq \left(\frac{4}{\pi}\right)^{n/2} \Gamma\left(\frac{n+2}{2}\right),$$

as well. Therefore

$$\kappa_n = \left(\frac{4}{\pi}\right)^{n/2} \Gamma\left(\frac{n+2}{2}\right).$$

□

5. Notes and Remarks

- The isoperimetric inequality

CHAPTER 4

Measures on Metric Spaces

Definition 0.0.20. *A function μ defined on the sets of a space Ω is called a* **measure** *on Ω if it satisfies the conditions*

(i) $\mu(E) \geq 0, \mu(E) \leq +\infty$ *for each* $E \subseteq \Omega$;
(ii) $\mu(\emptyset) = 0$;
(iii) $\mu(E) \leq \mu(F)$ *if* $E \subseteq F$;
(iv) *if* (E_n) *is a sequence of subsets of* Ω *then*

$$\mu(\cup_n E_n) \leq \sum_n \mu(E_n).$$

*If μ is a measure on the space Ω then $E \subseteq \Omega$ is said to be $\mu-$***measurable** *if whenever $A \subseteq E$ and $B \subseteq E^c$ then*

$$\mu(A \cup B) = \mu(A) + \mu(B).$$

Notice that $E \subseteq \Omega$ is $\mu-$measurable precisely when given $A \subseteq E$ and $B \subseteq E^c$,

$$\mu(A \cup B) \geq \mu(A) + \mu(B); \tag{0.1}$$

after all, (iv) assures that the reverse inequality holds regardless of E, A or B. In fact, we really need only show (0.1) for sets A, B of finite $\mu-$measure. (WHY?)

Theorem 0.0.21 (Caratheodory). *Let μ be a measure on the set Ω. Then*

(i) *if* $\mu(N) = 0$ *then* N *is* $\mu-$*measurable;*
(ii) E *is* $\mu-$*measurable if and only if* E^c *is;*
(iii) *if* (E_n) *is a sequence of* $\mu-$*measurable sets then* $\cup_n E_n$ *and* $\cap_n E_n$ *are too;*
(iv) *if* (E_n) *is a sequence of pairwise disjoint* $\mu-$*measurable sets then*

$$\mu(\cup_n E_n) = \sum_n \mu(E_n).$$

Proof. (of (i)) Suppose $A \subseteq N$ and $B \subseteq N^c$. Of course $\mu(A) = 0$. Hence

$$\mu(B) \leq \mu(A \cup B) \leq \mu(A) + \mu(B) = \mu(B).$$

Squeezy says all are equal and so N is $\mu-$measurable. (ii) is plain. To start the proof of (iii) and (iv), we'll show that the union of two $\mu-$measurable sets E, F is $\mu-$measurable. Take A and B to be sets of finite $\mu-$measure with

$$A \subseteq E \cup F, \quad B \subseteq (E \cup F)^c.$$

Notice

$$A \cup B = (A \cap E) \cup ((A \cup B) \cap E^c).$$

February 3, 2009

Since

$$A \cap E \subseteq E, \text{ and } (A \cup B) \cap E^c \subseteq E^c,$$

by E's measurability,

$$\mu(A \cup B) = \mu((A \cap E) \cup ((A \cup B) \cap E^c)) = \mu(A \cap E) + \mu((A \cup B) \cap E^c).$$

At the same time

$$(A \cup B) \cap E^c = (A \cap E^c) \cup B$$

with

$$A \cap E^c \subseteq F$$

and

$$B \subseteq (E \cup F)^c;$$

so F's $\mu-$measurability assures that

$$\mu((A \cup B) \cap E^c) = \mu(A \cap E^c) \cup B) = \mu(A \cap E^c) + \mu(B).$$

But E's measurability is still in effect and says that

$$\mu(A \cap E) + \mu(A \cap E^c) = \mu(A),$$

so

$$\begin{aligned} \mu(A \cup B) &= \mu(A \cap E) + \mu((A \cup B) \cap E^c) \\ &= \mu(A \cap E) + \mu(A \cap E^c) + \mu(B) \\ &= \mu(A) + \mu(B). \end{aligned}$$

Therefore if E and F are $\mu-$measurable then so is $E \cup F$. Next we consider a sequence (E_n) of pairwise disjoint $\mu-$measurable sets and write $E = \bigcup_n E_n$. Let A and B be sets (of finite $\mu-$measure) with

$$A \subseteq E, \quad B \subseteq E^c.$$

Each of $\cup_{m=1}^n E_m$ is $\mu-$measurable and

$$\begin{aligned} \mu(A \cup B) &\geq \mu\left([A \cap \cup_{m=1}^n E_m] \cup B\right) \\ &= \mu(A \cap (\cup_{m=1}^n E_m)) + \mu(B), \end{aligned}$$

because, after all,

$$A \cup (\cup_{m=1}^n E_m) \subseteq \cup_{m=1}^n E_m,$$

and

$$B \subseteq (\cup_n E_n)^c = \cap_n E_n^c \subseteq \cap_{m=1}^n (E_m^c) = (\cup_{m=1}^n E_m)^c.$$

Since the sets $E_n, E_{n-1}, \ldots, E_1$ are pairwise disjoint and $\mu-$measurable,

$$\begin{aligned} \mu(A \cap ((\cup_{m=1}^n E_m))) &= \mu([A \cap (\cup_{m=1}^{n-1} E_m)] \cup [A \cap E_n] \\ &= \mu(A \cap (\cup_{m=1}^{n-1} E_m)) + \mu(A \cap E_n) \\ &= \cdots = \mu(A \cap (\cup_{m=1}^{n-2} E_m)) + \mu(A \cap E_{n-1}) + \mu(A \cap E_n) \\ &= \cdots = \mu(A \cap E_1) + \mu(A \cap E_2) + \cdots + \mu(A \cap E_n). \end{aligned}$$

So we see that

$$\begin{aligned} \mu(A \cup B) &\geq \mu(A \cap (\cup_{m=1}^n E_n) + \mu(B) \\ &= \sum_{m=1}^n \mu(A \cap E_m) + \mu(B), \end{aligned}$$

and this is so for each n. Thus

$$\begin{aligned} \mu(A \cup B) &\geq \sum_n \mu(A \cap E_n) + \mu(B) \\ &\geq \mu(A \cap (\cup_n E_n)) + \mu(B) \quad \text{(since } \mu \text{ is subadditive)} \\ &= \mu(A \cup E) + \mu(B) \\ &= \mu(A) + \mu(B) \\ &\geq \mu(A \cup B). \end{aligned}$$

All are equal; in particular, $\mu(A \cup B)$ and $\mu(A) + \mu(B)$ are equal and $\cup_n E_n$ is $\mu-$measurable. Next if $B = \emptyset$ we get for any subset A of $\cup_n E_n$,

$$\mu(A) = \sum_n \mu(A \cap E_n).$$

Letting $A = \cup_n E_n$ gives

$$\mu(\cup_n E_n) = \sum_n \mu(E_n).$$

This gives (iv) and (iii), at least in the case the sets are pairwise disjoint.

Alas from (ii) and the fact that any finite union of $\mu-$measurable sets is $\mu-$measurable, we see that any finite intersection of $\mu-$measurable sets is also $\mu-$measurable .* To establish (iv) is now clear sailing: let (E_n) be a sequence of (not necessarily pairwise disjoint) $\mu-$measurable sets. Write $\cup_n E_n$ as follows:

$$\cup_n E_n = E_1 \cup (E_1 \backslash E_2) \cup (E_3 \backslash (E_1 \cup E_2)) \cup \cdots ;$$

the result is what we wanted - $\cup_n E_n$ is the union of a sequence of pairwise disjoint $\mu-$measurable sets and so is itself $\mu-$measurable. □

$\star$ It seems that this sentence belongs in the previous paragraph.

Definition 0.0.22. *A function τ defined on a class $\mathcal{C}$ of subsets of a space Ω is a* **premeasure** *if*

(i) $\emptyset \in \mathcal{C}$;
(ii) $0 \leq \tau(C) \leq \infty$ *for all* $C \in \mathcal{C}$;
(iii) $\tau(\emptyset) = 0$.

Theorem 0.0.23. *If τ is a premeasure defined on a class $\mathcal{C}$ of subsets of the space Ω then the set function μ*

$$\mu(E) := \inf \left\{ \sum_i \tau(C_i) : C_i \in \mathcal{C}, E \subseteq \bigcup_i C_i \right\}$$

is a measure.

Here, as usual, the infimum over an empty set is $+\infty$. The measure μ so constructed from τ is said to be constructed by **Method I.**

Proof. It is plain and easy to see that the set function μ as defined from τ satisfies properties (i)-(iii) of the definition of a measure. Only 'countable subadditivity' requires a bit more proof. So let (E_n) be a sequence of subsets of the space Ω, and let's show that

$$\mu(\cup_n E_n) \leq \sum_n \mu(E_n).$$

For sure we can assume $\sum_n \mu(E_n) < \infty$ and so, in particular, each $\mu(E_n) < \infty$ as well. Now for each n, choose a covering of E_n, say $\{C_j(n)\}_j$ from $\mathcal{C}$ so that for the always present $\epsilon > 0$ we have

$$\mu(E_n) \leq \sum_j \tau(C_j(n)) \leq \mu(E_n) + \frac{\epsilon}{2^n}.$$

Now $\{C_j(n) : j, n \in \mathbb{N}\}$ covers $\cup_n E_n$ and so

$$\begin{aligned} \mu(\cup_n E_n) &\leq \sum_{j,n} \tau(C_j(n)) \\ &\leq \sum_n \sum_j \tau(C_j(n)) \\ &\leq \sum_n \left(\mu(E_n) + \frac{\epsilon}{2^n}\right) \\ &= \sum_n \mu(E_n) + \epsilon. \end{aligned}$$

As expected, $\epsilon > 0$ was arbitrary and its arbitrariness leads us to the conclusion that

$$\mu(\cup_n E_n) \leq \sum_n \mu(E_n). \quad \square$$

Let τ be a premeasure on the family $\mathcal{C}$ of subsets of the **metric** space Ω. Let $\delta > 0$ and denote by

$$\mathcal{C}_\delta = \{C \in \mathcal{C} : \text{ diameter } C \leq \delta\}.$$

Denote by τ_δ the restriction of τ to C_δ. The result is a premeasure that generates a measure μ_δ on Ω by Method I. A look at μ_δ is worth taking:

$$\mu_\delta := \inf\left\{\sum_i \tau(C_i) : E \subseteq \cup_i C_i, C_i \in \mathcal{C}, \text{ diam } C_i \leq \delta\right\}.$$

Notice that μ_δ is a measure by Method I. It's plain that as δ gets smaller there are fewer sets with diameter $\leq \delta$ so $\mu_\delta(E)$ gets bigger. Hence we can define $\mu(E)$ by

$$\mu(E) := \sup_{\delta > 0} \mu_\delta(E),$$

and we say that τ generates μ by **Method II.**

Method I applies in a general setting, as opposed to Method II which relies on the metric structure present in the underlying structure set Ω.

THEOREM 0.0.24. *μ constructed by Method II is a measure.*

The only possible stumbling point to this is the countable subadditivity so let's see why μ is countably subadditive. To this end, let (E_n) be a sequence of subsets of Ω and consider these quantities

$$\mu(\cup_n E_n) \text{ and } \sum_n \mu(E_n).$$

Obviously the latter exceeds the former if it's ∞ so we may as well assume $\sum_n \mu(E_n) < \infty$.

Now for each $\delta > 0$, μ_δ is a known measure so

$$\mu_\delta(\cup_n E_n) \leq \sum_n \mu_\delta(E_n),$$

which in turn is

$$\leq \sum_n \mu(E_n).$$

The countable subadditivity follows from this. □

A key ingredient to our mix is provided in the following. Comment on this theorem?

THEOREM 0.0.25. *Let μ be a measure on a metric space Ω derived from the premeasure τ by Method II. If A and B are non-empty subsets of Ω that are 'positively separated' then*

$$\mu(A \cup B) = \mu(A) + \mu(B).$$

Here A and B 'positively separated' means that there is a $\delta > 0$ so that for any $a \in A$ and $b \in B$,

$$\rho(a, b) \geq \delta.$$

Measures that enjoy this additive property are called **metric measures** and their importance lies in the fact that these are precisely the measures on a metric space for which every Borel set is measurable. Define Borel sets here or somewhere else?

PROOF. We need to show that if A and B are positively separated then

$$\mu(A \cup B) \geq \mu(A) + \mu(B)$$

where all the terms involved are finite. The idea of the proof is to cover A, B and $A \cup B$ with very fine covers from the domain $\mathcal{C}$ of τ, so fine that we can distinguish between which sets are needed to cover A and which to cover B.

More precisely, suppose

$$\rho(a, b) \geq \delta > 0$$

for any $a \in A$ and $b \in B$. Let $\epsilon > 0$ announce its presence. By the rules of engagement

$$\mu(A \cup B) = \sup_{d>0} \inf \left\{ \sum_i \tau(C_i) : \operatorname{diam} C_i \leq d, C_i \in \mathcal{C}, A \cup B \subseteq \bigcup_i C_i \right\}.$$

So we can choose a sequence (C_i) in $\mathcal{C}$ so that

- diam C_i are all 'small';
- $A \cup B \subseteq \cup_i C_i$;
- $\sum \tau(C_i) \leq \mu(A \cup B) + \epsilon$.

'Small'? Well to set things up for computation of the $\mu-$measure of A and of B through the intermediaries $\mu_\delta(A), \mu_\delta(B)$, we let δ_1, δ_2 be positive numbers each less than diam Ω, and let η be the number

$$\eta = \min \left\{ \delta_1, \delta_2, \frac{\delta}{3} \right\}.$$

Choose η as our model of small.

Here's the first punchline; **if each C_i has diameter $\leq \eta$ then each has diameter $\leq \frac{\delta}{3}$ and so a given C_i can intersect A or B but not both.** It follows that

$$\sum_{C_i \cap A \neq \emptyset} \tau(C_i) + \sum_{C_i \cap B \neq \emptyset} \tau(B_i) \leq \sum_i \tau(C_i) \leq \mu(A \cup B) + \epsilon.$$

But each C_i that intersects A has diameter $\leq \eta \leq \delta_1$ so knowing, as we do, that A is a subset of

$$\bigcup_{C_i \cap A \neq \emptyset} C_i,$$

we get

$$\mu_{\delta_1}(A) \leq \sum_{C_i \cap A \neq \emptyset} \tau(C_i).$$

Similarly,

$$\mu_{\delta_2}(B) \leq \sum_{C_i \cap B \neq \emptyset} \tau(C_i).$$

The result:

$$\mu_{\delta_1}(A) + \mu_{\delta_2}(B) \leq \sum_{C_i \cap A \neq \emptyset} \tau(C_i) + \sum_{C_i \cap B \neq \emptyset} \tau(C_i) \leq \mu(A \cup B) + \epsilon,$$

and so Method II leads us to believe

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

Since $\epsilon > 0$ was arbitrary, this proof is done. □

Metric measures enjoy some very strong continuity properties. Here's one of them. Comment on the importance of this proposition.

PROPOSITION 0.0.26. *Let μ be a metric measure on a metric space Ω. Suppose (A_n) is an increasing sequence of subsets of Ω so that A_n and A_{n+1}^c are positively separated. Then*

$$\mu(\cup_n A_n) = \sup_n \mu(A_n).$$

We defer the proof until after giving the Proposition a chance to 'show off.'

THEOREM 0.0.27. *If μ is a metric measure on the metric space Ω then every closed subset of Ω is $\mu-$measurable.*

Comment that this tells us that Borel sets are $\mu-$measurable.

PROOF. Suppose F is a closed subset of the metric space Ω, and let $A \subseteq F$ and $B \subseteq F^c$ be non-empty sets. For each n let

$$B_n := \left\{ x \in B : \inf_{y \in F} \rho(x, y) > \frac{1}{n} \right\}.$$

Notice that $B_n \nearrow B$.

By their very definition, each B_n is positively separated from A. We'll show that each B_n is also positively separated from B_{n+1}^c. To this end, let $x \in B_n$ and $u \in B_{n+1}^c$. Now $u \notin B_{n+1}$ so

$$\frac{1}{n+1} \geq \inf_{y \in F} \rho(x, y);$$

There is a $y_0 \in F$ so that

$$\frac{1}{n+\frac{1}{2}} \geq \rho(u, y_0).$$

Now if

$$\rho(x,u) \leq \frac{\frac{1}{2}}{n\left(n+\frac{1}{2}\right)},$$

then

$$\begin{aligned} \inf_{y\in F} \rho(x,y) &\leq \rho(x,y_0) \\ &\leq \rho(x,u) + \rho(u,y_0) \\ &\leq \frac{\frac{1}{2}}{n(n+\frac{1}{2})} + \frac{1}{\left(n+\frac{1}{2}\right)}, \end{aligned}$$

which by the method of common denominators, is

$$= \frac{\frac{1}{2}+n}{n\left(n+\frac{1}{2}\right)} = \frac{1}{n}.$$

Hence $x \notin B_n$. Our conclusion? If $x \in B_n$ and $u \in B_{n+1}^c$, then

$$\rho(x,u) > \frac{\frac{1}{2}}{n\left(n+\frac{1}{2}\right)} = \frac{1}{n(2n+1)}.$$

i.e., B_n and B_{n+1}^c are positively separated.

Let's compute $\mu(A \cup B)$:

$$\begin{aligned} \mu(A\cup B) &\geq \sup_n \mu(A\cup B_n) \\ &= \sup_n \mu(A) + \mu(B_n) \\ &= \mu(A) + \sup_n \mu(B_n) \\ &= \mu(A) + \mu(\cup_n B_n) \text{ by Proposition 0.0.26} \\ &= \mu(A) + \mu(B), \end{aligned}$$

and F is measurable. □

Now we look to the rather elegant proof of Proposition 0.0.26:

Proof. We have

$$A_1 \subseteq A_2 \subseteq \cdots \subseteq A_n \subseteq \cdots$$

with A_n and A_{n+1}^c positively separated for each n. We want to show that $\mu(\cup_n A_n) \leq \sup_n \mu(A_n)$ and in this effort we may plainly suppose $\sup_n \mu(A_n) < \infty$ since otherwise all is okay.

First we look at the difference sequence

$$D_1 = A_1, D_2 = A_2 \backslash A_1, D_3 = A_3 \backslash A_2, \ldots, D_n = A_n \backslash A_{n-1}, \ldots$$

Of course, $D_n \subseteq A_n$ but further

$$D_2 \subseteq A_1^c, D_3 \subseteq A_2^c \subseteq A_1^c, D_4 \subseteq A_3^c \subseteq A_2^c \subseteq A_1^c, \ldots$$

So

$$D_1 \subseteq A_1 \text{ and } D_3, D_4, \ldots \subseteq A_2^c,$$
$$D_2 \subseteq A_2 \text{ and } D_4, D_5, \ldots \subseteq A_3^c,$$

etc., etc., etc. In particular

$$\begin{array}{rl} D_1 \subseteq A_1, & D_3 \subseteq A_2^c \\ D_2 \subseteq A_2, & D_4 \subseteq A_3^c \\ D_1 \cup D_3 \subseteq A_3, & D_5 \subseteq A_4^c \\ D_2 \cup D_4 \subseteq A_4, & D_6 \subseteq A_5^c \\ & \vdots \\ D_1 \cup D_3 \cup \cdots \cup D_{2n-1} \subseteq A_{2n-1}, & D_{2n+1} \subseteq A_{2n}^c \\ D_2 \cup D_4 \cup \cdots \cup D_{2n} \subseteq A_{2n}, & D_{2n} \subseteq A_{2n+1}^c. \end{array}$$

It follows (inductively if you must know) that

$$\mu(D_1) + \mu(D_3) + \cdots + \mu(D_{2n-1}) + \mu(D_{2n+1}) = \mu(\cup_{k=1}^{n+1} D_{2k-1}),$$

and

$$\mu(D_2) + \mu(D_4) + \cdots + \mu(D_{2n}) + \mu(D_{2n+2}) = \mu(\cup_{k=1}^{n+1} D_{2k}).$$

Each of $\cup_{k=1}^{n+1} D_{2k-1}$ and $\cup_{k=1}^{n+1} D_{2k}$ are subsets of A_{2n+2} and so all above find themselves

$$\leq \mu(A_{2n+2}) \leq \sup_n \mu(A_n) < \infty.$$

Conclusion: both series $\sum_n \mu(D_{2n-1})$ and $\sum_n \mu(D_{2n})$ converge. Now

$$\begin{aligned} \mu(\cup_n A_n) &= \mu(A_n \cup D_{n+1} \cup D_{n+2} \cup \cdots) \quad \text{Not sure about this one.} \\ &\leq \mu(A_n) + \mu(D_{n+1}) + \mu(D_{n+2}) + \cdots \\ &\leq \sup_n \mu(A_n) + \sum_{k=n+1}^{\infty} \mu(D_k); \end{aligned}$$

if $\epsilon > 0$ be given then there is an n so that the latter sum is less than ϵ so

$$\mu(\cup_n A_n) \leq \sup_n \mu(A_n) + \epsilon.$$

Enough said. □

We've seen that every Borel set in a metric space Ω is $\mu-$measurable whenever μ is a metric measure on Ω. It's worthwhile to note that this is **not** accidental; it's part and parcel of being a metric measure. Indeed if we suppose μ is a measure on the metric space Ω for which every closed subset of Ω is $\mu-$measurable and if $A, B \subseteq \Omega$ are positively separated (so $\overline{A} \cap \overline{B} = \emptyset$), then

$$\mu(A \cup B) = \mu((A \cup B) \cap \overline{A}) + \mu((A \cup B) \cap \overline{A}^c)$$

by the measurability of $\overline{A}$. But

$$(A \cup B) \cap \overline{A} = A \text{ and } (A \cup B) \cap (\overline{A})^c = B$$

by the positive separation of A and B. So

$$\mu(A \cup B) = \mu((A \cup B) \cap \overline{A}) + \mu((A \cup B) \cap (\overline{A})^c = \mu(A) + \mu(B),$$

and μ is a metric measure.

We turn now to regularity properties of measures on a metric space, that is, how well we can approximate a typical value $\mu(E)$ of μ by values at 'good' sets - closed, open, $\mathcal{F}_\sigma, \mathcal{G}_\delta$ or even compact. REWRITE THIS SENTENCE: Before we go any further we hasten to say that the problems we're concerned with are previously with metric measures that are **not** finite. To be sure, we emphasize that if measure is finite then good things happen (at least in the world of Borel sets).

Here's (some of) what's so.

THEOREM 0.0.28. *Let (X, ρ) be a metric space. Let μ be a countably additive map from the Borel $\sigma-$field $\mathcal{B}\wr(X)$ to $[0, \infty)$. Define this $\sigma-$field?*

(i) *For any Borel set B in X, we have*

$$\sup\{\mu(F) : F \subseteq B, F \text{ closed}\} = \mu(B) = \inf\{\mu(G) : B \subseteq G, G \text{ open}\}.$$

Should we change G to U since we use U in the proof?

(ii) *If X is Polish (that is, there is a complete metric that generates X's topology <u>and</u> X is separable) then for any Borel set B in X*

$$\mu(B) = \sup\{\mu(K) : K \subseteq B, K \text{ compact}\}.$$

This result finds so much use in what we do, we submit a proof in detail.

PROOF. (of (i)) Here's the 'trick' to find an approach to general Borel sets; we consider the collection $\mathcal{S}$ of Borel sets in X that satisfy the conclusion of (i).

Open sets belong to $\mathcal{S}$. If U is open then it's trivial to approximate $\mu(U)$ from above by μ's values on open super sets of U. How to approximate from within by μ's values at closed subsets? Well, notice that if U is an open set in X then U^c is closed and so $x \in U^c$ precisely when $d(x, U^c) = 0$ where

$$d(x, U^c) = \inf\{d(x, y) : y \in U^c\}.$$

But $d(x, U^c)$ is a continuous function of $x \in X$ and so we see that

$$U = \bigcup_{n \in \mathbb{N}} \left[d(\cdot, U^c) \geq \frac{1}{n} \right],$$

and U is an $\mathcal{F}_\sigma-$set. If we let $F_n = \left[d(\cdot, U^c) \geq \frac{1}{n}\right]$ then each F_n is closed,

$$F_1 \subseteq F_2 \subseteq \cdots \subseteq F_n \subseteq \cdots \nearrow U,$$

and so

$$\mu(U) = \lim_n \mu(F_n).$$

Open sets are $\mathcal{F}_\sigma$'s and so each open set belongs to $\mathcal{S}$.

Now $\mathcal{S}$, by it's very definition, can be described in a manner suited to epsilonics. Indeed, a Borel set $B \in \mathcal{S}$ precisely when B satisfies

given $\epsilon > 0$ there is an open set U containing B and a closed set F contained in B such that $\mu(U \backslash F) < \epsilon$.

It's plain from this that $\mathcal{S}$ is closed under complements.

What about countable unions? We'll suppose (B_k) is a sequence of Borel sets each one of which is in $\mathcal{S}$. For each n, pick U_n and F_n so that U_n is open, F_n is closed, $F_n \subseteq B_n \subseteq U_n$ and

$$\mu(U_n \backslash F_n) < \frac{\epsilon}{2^{n+1}}.$$

Let $U = \cup_n U_n, F = \cup_n F_n$. Then

$$F \subseteq \cup_n B_n \subseteq U.$$

Moreover

$$\begin{aligned} \mu(U \backslash F) &= \mu(\cup_n U_n \backslash \cup_n F_n) \leq \mu(\cup_n (U_n \backslash F_n)) \\ &\leq \sum_n \mu(U_n \backslash F_n) \leq \sum_n \frac{\epsilon}{2^{n+1}} = \frac{\epsilon}{2}. \end{aligned}$$

Now U is open and F is almost closed. In fact, $F = \cup_n F_n$ so there is an n_0 so that

$$\mu(F \backslash \cup_{n>n_0} F_n) < \frac{\epsilon}{2};$$

The result:

$$\bigcup_{n=1}^{n_0} F_n \subseteq \bigcup_n B_n \subseteq U,$$

where $\bigcup_{n=1}^{n_0} F_n$ is closed, U is still open and

$$\mu(U \backslash \cup_{n=1}^{n_0} F_n) < \epsilon.$$

Therefore $\mathcal{S}$ is a collection of Borel sets containing X's topology, (containing?) closed under complementation and the taking of countable unions. Thus $\mathcal{S} = \mathcal{B}_0(X)$.

(ii) hints at the plentitude of compact sets in Polish spaces. This comes from the (existence of a) complete metric; in complete metric spaces, subsets are relatively compact precisely when they're totally bounded. So to prove (ii) we'll use the separability to construct good closed and totally bounded approximants. We start with a complete separable metric space (X, ρ) and show (ii) holds for the Borel set X itself. Let $\{x_n : n \in \mathbb{N}\}$ be a countable dense subset of X. For each $n \in \mathbb{N}$

$$X = \cup_k B_{1/n}(x_k)$$

where

$$B_\epsilon(x) = \{y \in X : \rho(x, y) < \epsilon\}.$$

For each n we can find a k_n so that

$$\mu\left(\bigcup_{k=1}^{k_n} B_{1/n}(x_k)\right)$$

is within $\frac{\epsilon}{2^n}$ of $\mu(X)$. Look at

$$K = \bigcap_n \left(\bigcup_{k=1}^{k_n} \overline{B_{1/n}(x_k)}\right).$$

Since K is complete, it's closed. K is totally bounded, too: after all, for any n,

$$K \subseteq \bigcup_{k=1}^{k_n} \overline{B_{1/n}(x_k)}$$

so each point of K is within $2/n$ of an $x_k, k = 1, \ldots, k_n$. Therefore K is compact.

$$\begin{aligned}
\mu(K^c) &= \mu\left(\left(\bigcap_n \left[\bigcup_{k=1}^{k_n} \overline{B_{1/n}(x_k)}\right]\right)^c\right) \\
&= \mu\left(\bigcup_n \left[\bigcup_{k=1}^{k_n} \overline{B_{1/n}(x_k)}\right]^c\right) \\
&\leq \sum_n \mu\left(\left[\bigcup_{k=1}^{k_n} \overline{B_{1/n}(x_k)}\right]^c\right) \\
&\leq \sum_n \mu\left(\left[\bigcup_{k=1}^{k_n} B_{1/n}(x_k)\right]^c\right) \\
&\leq \sum_n \frac{\epsilon}{2^n} = \epsilon.
\end{aligned}$$

Generally if we apply (i) we can find inside any Borel subset B of X a closed set F with $\mu(F)$ within $\epsilon/2$ of $\mu(B)$; then define $\tilde{\mu}$ on the Borel sets of X by

$$\tilde{\mu}(E) = \mu(E \cap F).$$

Our opening salvo applied to $\tilde{\mu}$ gives a compact $K \subseteq X$ so $\tilde{\mu}(K)$ is within $\epsilon/2$ of $\tilde{\mu}(X) = \mu(F)$. Then $K \cap F$ does the dirty deed since $\mu(K \cap F) = \tilde{\mu}(K)$, which is within ϵ of $\mu(B)$. □

Before entering into the subject of regularity of metric measures, there are a few apt comments to be made regarding the theorem above. First regarding (ii) an example might highlight the completeness hypothesis. Suppose Ω is a subset of $[0,1]$ for which the inner measure $\lambda_k(\Omega)$ is 0 and the outer measure $\lambda_k(\Omega)$ is 1. Of course Ω is not Lebesgue measurable and hopelessly incomplete. However Ω is separable. Now the Borel subsets of Ω are just those subsets of Ω of the form $B \cap \Omega$ where B is a Borel subset of $[0,1]$. If we define $P(B \cap \Omega)$ to be the Lebesgue measure of B then P is $\lambda^*|_{B_0}(\Omega)$. P is a probability Borel measure on the separable metric space Ω. However, if K is a compact subset of Ω then K is also a compact subset of $[0,1]$ and so a Borel subset of $[0,1]$; it follows that

$$P(K) = \lambda^*(K) = \lambda_*(K) = 0.$$

P is not a regular Borel measure on Ω. Comments on this example:

- Should we say something about inner measure and outer measure? What is λ_k?
- How do we know anything about Ω? eg Why is Ω not Lebesgue measurable and incomplete?
- What is λ^*?
- Should we define probability here or somewhere prior to this?

Some kind of completeness assumption is needed in (ii). How about separability? Well most probabilities have separable support. Should we define 'separable support?' Here's the fact.

THEOREM 0.0.29. *Let Ω be a metric space. Then in order that every probability Borel measure on Ω have a separable support it is both necessary and sufficient that each discrete subset of Ω have 'non-measurable cardinal.'*

Recall that a set S has *measurable cardinal* if one can define on 2^S a probability measure that vanishes on singletons; otherwise S has *non-measurable cardinal*. It's unknown if sets measurable cardinals exist. It is known that if they do, then they must be huge! So the comforts of finite measure aside, let's discuss general metric measures and approximation properties thereof. First, a general definition: if μ is a measure on a space Ω and $\mathcal{R}$ is a family of subsets of Ω then we call μ **$\mathcal{R}$-regular** if for any $E \subseteq \Omega$ there is an $R \in \mathcal{R}$ so that $E \subseteq R$ and $\mu(E) = \mu(\mathcal{R})$.

Now some facts.

- If μ is a measure on a space Ω generated by a premeasure τ on a collection $\mathcal{C}$ with $\Omega \in \mathcal{C}$ by Method I, then μ is $C_{\sigma\delta}-$regular. What does this mean? Since $\Omega \in \mathcal{C}, \Omega$ is certainly $C_{\sigma\delta}$ and so for sets E with $\mu(E) = \infty, E \subseteq \Omega$ serves us well. So we look at E's with $\mu(E) < \infty$. In this case, epsilonics enter the foray: for each n there is a sequence $(C_i^n)_i$ of members of $\mathcal{C}$ such that
$$E \subseteq \cup_i C_i^n, \text{ and } \mu(E) \leq \sum_i \tau(C_i^n) \leq \mu(E) + \frac{1}{n}.$$
Set
$$D = \cap_n \cup_i C_i^n \in C_{\sigma\delta}.$$
Then $E \subseteq D$ and for each n
$$\mu(E) \leq \mu(D) \leq \sum_i \tau(C_i^n) \leq \mu(E) + \frac{1}{n}.$$
It follows that $\mu(E) = \mu(D)$.
- Let μ be a measure on a metric space Ω generated by a premeasure τ on a class $\mathcal{C}$ which contains Ω via Method II. Then μ is $C_{\sigma\delta}-$regular.

  PROOF. As before we use $\Omega \in \mathcal{C}$ to rid us nuisances: take $E \subseteq \Omega$; if $\mu(E) = \infty$ then $\Omega \in \mathcal{C} \subseteq C_{\sigma\delta}$ contains E and plainly $\mu(E) = \mu(\Omega) = \infty$. So we can restrict our attention to E such that $\mu(E) < \infty$. But now each $\mu_\eta(E)$ is also finite since $\mu_\eta \leq \mu$ for any $\eta > 0$. Because each $\mu_\eta(E) < \infty$ we can also find a $C_\eta \in C_{\sigma\delta}$ such that $E \subseteq C_\eta$ and $\mu_\eta(E) = \mu_\eta(C_n)$. To keep tabs on $\mu(E)$ for each n let $C_n \in C_{\sigma\delta}$ be so that $E \subseteq C_n$ and
$$\mu_{1/n}(E) = \mu_{1/n}(C_n).$$
Put
$$C = \cap_n C_n \in C_{\sigma\delta}.$$
$E \subseteq C$ and
$$\mu_{1/n}(E) \leq \mu_{1/n}(C) \leq \mu_{1/n}(C_n) = \mu_{1/n}(E) \leq \mu(E).$$
Choosing carefully from this we see
$$\mu_{1/n}(E) \leq \mu_{1/n}(C) \leq \mu(E).$$
Taking $n \to \infty$, we get
$$\mu(E) = \mu(C),$$
and that's all she wrote.

Each of these facts has interesting consequences in case $\mathcal{C}$ is special.

COROLLARY 0.0.30. *Suppose μ is a measure on a topological group generated by a premeasure τ defined on the topology of the space by Method I. Then μ is $\mathcal{G}_\delta-$regular. If μ is a measure on a metric space generated by a premeasure defined on the topology of the space by Method II, then μ is $\mathcal{G}_\delta-$regular.*

COROLLARY 0.0.31. *Suppose μ is a measure on a topological space Ω generated by a premeasure defined on the Borel subsets of Ω by Method I. Then μ is Borel-regular. Suppose μ is a measure on a metric space Ω generated by a premeasure defined on the Borel subsets of Ω by Method II. then μ is Borel-regular.*

We have the following general principle for approximating from within:

If μ is an $\mathcal{R}$-regular measure on a space Ω and E is a $\mu-$measurable subset of Ω with $\mu(E) < \infty$ then there is a set $R_1 \backslash R_2$ where $R_1, R_2 \in \mathcal{R}$, contained in E so that

$$\mu(R_1 \backslash R_2) = \mu(E).$$

Choose $R_1 \in \mathcal{R}$ with $E \subseteq R_1$ so that

$$\mu(E) = \mu(R_1).$$

E is $\mu-$measurable so

$$\mu(E0 = \mu(R_1) = \mu(R_1 \cap E) + \mu(R_1 \cap E^c) = \mu(E) + \mu(R_1 \cap E^c).$$

But $\mu(E) < \infty$ so

$$\mu(R_1 \cap E^c) = 0.$$

μ is <u>still</u> $\mathcal{R}-$regular so there is an $R_2 \in \mathcal{R}$ so that

$$R_1 \cap E^c \subseteq R_2 \text{ and } 0 = \mu(R_1 \cap E^c) = \mu(R_2).$$

Now

$$\begin{aligned} R_1 \backslash R_2 &\subseteq (R_1 \cap E^c) \\ &= R_1 \cap (R_1 \cap E6c)^c \\ &= R - 1 \cap (R_1^c \cup E) \\ &= (R_1 \cap R_1^c) \cup (R_1 \cap E) = E. \end{aligned}$$

Also $\mu(R_2) = 0$. So

$$\begin{aligned} \mu(R_1 \backslash R_2) &= \mu(R_1 \backslash R_2) + \mu(R_2) \\ &\geq \mu((R_1 \backslash R_2) \cup R_2) \\ &\geq \mu(R_1) \\ &= \mu(E) \\ &\geq \mu(R_1 \backslash R_2). \end{aligned}$$

Voila!

CHAPTER 5

Banach and Measure

Like most (abstract) analysts of his day, Banach also took a keen interest in developments related to the existence, uniqueness and uses of Haar measure. In this chapter we recount Banach's views on the subject. Naturally we open with a bit of functional analysis and present the real case of the Hahn-Banach theorem. We follow this with a derivation of the existence of "Banach limits." After a brief discussion of weak topologies, we pass to Banach's remarkable characterization of weakly null sequences of bounded functions on a set. That is followed by Banach's characterization of weakly null sequences in the space $C(Q)$, where Q is a compact metric space, a result that was derived *before* Banach knew what $C(Q)^*$ was! It wasn't long before Banach was able to 'compute' $C(Q)^*$. We follow his way of doing this, benefiting from his view of abstract Lebesgue integration, à la a Daniell type construction. Finally we present Banach's construction of invariant measures. His proof was a source of inspiration for many and will reappear in our discussion of Steinlage's description of existence and uniqueness of measures on locally compact spaces that are invariant under group actions.

1. A bit of Functional Analysis

Theorem 1.0.32 (Hahn-Banach Theorem). *Let X be a* real *linear space, and let S be a linear subspace of X. Suppose that $p : X \to \mathbb{R}$ is a subadditive positively homogeneous functional and $f : S \to \mathbb{R}$ is a linear functional with $f(s) \leq p(s)$ for all $s \in S$. Then there is a linear functional F defined on all of X such that $F(x) \leq p(x)$ for all $x \in X$, and $F(s) = f(s)$ for all $s \in S$.*

Proof. Our first task is to see how to extend a functional like f one dimension at a time preserving the domination by p. With this in mind, let $x \in X \backslash S$ and notice that for any linear combination $s + \alpha x$ of a vector in S and x, whatever the linear extension F's value at x (say that value is 'c') we must have

$$F(s + \alpha x) = F(s) + \alpha F(x) = f(s) + \alpha c.$$

So it must be (if we are to have F dominated by p) that

$$f(s) + \alpha c \leq p(s + \alpha x)$$

holds for all $s \in S$ and all real α's. We'll follow where this leads us; for all $\alpha \in \mathbb{R}$ and $s \in S$ we have to have

$$\alpha c \leq p(s + \alpha x) - f(s).$$

For $\alpha > 0$ this tells us that

$$\begin{aligned} c &\leq \frac{1}{\alpha}\left(p(s + \alpha x) - f(s)\right) \\ &= p\left(\frac{s}{\alpha} + x\right) - f\left(\frac{s}{\alpha}\right); \end{aligned}$$

while if $\alpha < 0$ then ($-\alpha > 0$ and)

$$-\alpha c \geq f(s) - p(s + \alpha x),$$

or since $-\alpha > 0$,

$$\begin{aligned} c &\geq -\frac{1}{\alpha} f(s) - \frac{1}{-\alpha} p(s + \alpha x) \\ &= f\left(\frac{s}{-\alpha}\right) - p\left(\frac{s}{-\alpha} - x\right) \end{aligned}$$

Taking into account the linearity of S we see that what we seek is a $c \in \mathbb{R}$ (which will be $F(x)$) so that regardless of $s, s^{'} \in S$ satisfies

$$f(s^{'}) - p(s^{'} - x) \leq c \leq p(s + x) - f(s).$$

Does such a c exist? You bet! After all if $s, s^{'} \in S$ then

$$f(s) + f(s^{'}) = f(s + s^{'}) \leq p(s + s^{'}) \leq p(s - x) + p(s^{'} + x).$$

So for all $s, s^{'} \in S$

$$f(s) - p(s - x) \leq p(s^{'} + x) - f(s^{'}),$$

and we *can* chose c in an appropriate manner.

Now we know that we can extend linear functionals one dimension at a time while preserving p's domination. It's time for some transfinite hijinks. We consider the collection of all linear functionals g defined on a linear subspace Y of X such that $S \subseteq Y$, $g|_S = f$ and on Y, $g \leq p$. We partially order this collection by saying that "$g_1 \leq g_2$" if g_2 is an extension of g_1 (so g_1 is defined on a linear subspace Y_1 that is contained in g_2's domain).

Hausdorff's maximal principle ensures us that there is a maximal linearly ordered subfamily $\{g_\alpha\}$ of linear extensions of f so that on g_α's domain Y_α (which contains S), $g_\alpha \leq p$. We define F on the linear space that's the union of the domains of the g_α's as one mine expect! $F(x) = g_\alpha(x)$ if x is in g_α's domain. Because of the ordering described above, the domain of F is linear subspace of X and on that domain F is linear and dominated by p. Of course, F is a linear extension of f and the domain of F *must be all of* X - this is assured us by the opening salvo. □

We put the Hahn-Banach theorem to immediate use by establishing the existence of 'generalized limits' or Banach limits, as we'll refer to them henceforth. We will call on two spaces: l^∞, the space of bounded, real-valued sequences, and the linear subspace c of l^∞ consisting of all the convergent sequences. Typically if $x \in l^\infty$ then

$$||x||_\infty = \sup\{(x_n) : x \in \mathbb{N}\}.$$

THEOREM 1.0.33 (Banach). *There exists a linear functional LIM on l^∞ such that*

(i) *LIM*$(x) \geq 0$ *if* $x = (x_n) \in l^\infty$ *and* $x_n \geq 0$ *for all* n;
(ii) $|LIM(x)| \leq ||x||_\infty$, *for all* $x \in l^\infty$;
(iii) *If* $x \in l^\infty$ *and* $Tx = (x_2, x_3, \dots)$ *for* $x = (x_1, x_2, \dots)$, *then* $LIM(Tx) = LIM(x)$;
(iv) *For any* $x \in l^\infty$

$$\liminf_{n\to\infty} x_n \leq LIM(x) \leq \limsup_{n\to\infty} x_n.$$

PROOF. Let $p : l^\infty \to \mathbb{R}$ be given by

$$p(x) = \limsup_{n\to\infty} \frac{x_1 + \cdots + x_n}{n}.$$

It is easy to see that p is subadditive and positively homogeneous. Next, let $f : c \to \mathbb{R}$ be the linear functional

$$f(x) = \lim_n x_n.$$

Since f and p agree on c, $f(x) \le p(x)$ is trivially satisfied. LIM is any Hahn-Banach extension of f to all of l^∞. LIM is a linear functional on l^∞ such that for any $x \in l^\infty$

$$\mathrm{LIM}(x) \le p(x).$$

To see (i), take $x \in l^\infty$ and suppose $x_n \ge 0$ for all n. Then

$$-\mathrm{LIM}(x) = \mathrm{LIM}(-x) \le p(-x),$$

so

$$\begin{aligned} \mathrm{LIM}(x) &\ge -p(x) = -\limsup_{n\to\infty} \frac{-x_1 - x_2 - \cdots - x_n}{n} \\ &= \liminf_{n\to\infty} \frac{x_1 + x_2 + \cdots + x_n}{n} \ge 0, \end{aligned}$$

because $x_n \ge 0$ for all n.

(ii) is plain since $p(x) \le ||x||_\infty$ for all $x \in l^\infty$. To see (iii), take $x \in l^\infty$, then

$$p(x - Tx) = \limsup_{n\to\infty} \frac{x_1 - x_{n+1}}{n} = 0,$$

since x *is* bounded. It follows that

$$\mathrm{LIM}(x - Tx) \le p(x - Tx) = 0.$$

So for any $x \in l^\infty$

$$\mathrm{LIM}(x) \le \mathrm{LIM}(Tx). \tag{1.1}$$

This applies as well to $-x$ so

$$-\mathrm{LIM}(x) = \mathrm{LIM}(-x) \le \mathrm{LIM}(-Tx) = -\mathrm{LIM}(Tx),$$

and so

$$\mathrm{LIM}(x) \ge \mathrm{LIM}(Tx) \tag{1.2}$$

as well. $\mathrm{LIM}(x) = \mathrm{LIM}(Tx)$ is the only conclusion that can be drawn from (1.1) and (1.2).

For (iv), let $\epsilon > 0$. Find $N \in \mathbb{N}$ so

$$\inf_n x_n \le x_N \le \inf_n x_n + \epsilon.$$

Then

$$x_n + \epsilon - x_N \ge \inf_n x_n + \epsilon - x_N \ge 0.$$

Hence

$$0 \le \mathrm{LIM}(x + \epsilon - x_N) = \mathrm{LIM}(x) + \epsilon - x_N$$

by (i). Hence

$$\inf_n x_n \le x_N \le \mathrm{LIM}(x) + \epsilon.$$

Since $\epsilon > 0$ was arbitrary

$$\inf_n x_n \le \mathrm{LIM}(x).$$

For any k

$$\inf_{n\ge k} x_n = \inf_{n\ge k} T^k x_n \le \mathrm{LIM}(T^k x) = \mathrm{LIM}(x);$$

but

$$\liminf_n x_n = \sup_n \inf_{k\ge n} x_k \le \sup_n \mathrm{LIM}(x) = \mathrm{LIM}(x).$$

Again using LIM's linearity, we see that

$$\limsup_n x_n = -\liminf_{n\longrightarrow\infty}(-x_n) \ge -\mathrm{LIM}(-x) = \mathrm{LIM}(x).$$

The proof of the existence of Banach limits appears. □

Let X be a normed linear space. A linear functional on X is **bounded** if there is a $c > 0$ so that $|f(x) \le c||x||$ for all $x \in X$. The appellation 'bounded' pertains to f's boundedness on the closed unit ball

$$B_X = \{x \in X : ||x|| \le 1\}$$

of X. It is easy to show that a linear functional's boundedness is tantamount to its continuity and this, in turn, is assured by f's continuity at the origin. Then

$$||f|| = \sup_{x\in B_X} |f(x)|$$

is a norm on X^* and is in fact, 'complete.' So Cauchy sequences in X^* with respect to this just-defined norm, converge.

The Hahn-Banach Theorem provides any normed linear space with lots of linear functionals in its dual.

- Let S be a linear subspace of the real normed linear space X, and let s^* be a continuous linear functional on S. Then there is an extension x^* of s^* to all of X so that $||s^*|| = ||x^*||$. We simply let $p(x) = ||s^*||\,||x^*||$ and apply the Hahn-Banach Theorem to find a linear functional F defined on all of X with $F(x) \le p(x)$ for all $x \in X$. Because

  $$F(-x) \le p(-x) = p(x),$$

  we see that

  $$F(x) \le p(x) = ||s^*||\,||x^*||$$

  for all $x \in X$ and so $F = x^*$ *is* in X^* and $||F|| \le ||s^*||$. Since F extends s^*, $||s^*|| \le ||F||$ too.
- Let x be a member of the (real) normed linear space X. Then there is an $x^* \in X^*$ with $||x^*|| = 1$ so that $x^*(x) = ||x||$. Indeed, let

  $$S = \{\alpha x : \alpha \in \mathbb{R}\},$$

  define $s^*(\alpha x) = \alpha||x||$, apply the Hahn-Banach Theorem and get x^*, the norm-preserving extension of s^*.

It is frequently the case that for a particular normed linear space the norm-topology is too coarse to uncover special and delicate phenomena peculiar to that space. Oft times such phenomena are best described using the weak topology.

Definition 1.0.34. *A base for the* **weak topology** *of the normed linear space X is given by sets of the form*

$$U(x; x_1^*, x_2^*, \dots, x_n^*) = \{y \in X : |x_i^*(x - y)| < \epsilon, i = 1, \dots n\}$$

where $x \in X, x_1^*, \dots, x_n^* \in X^*$, and $\epsilon > 0$.

The topology generated by this base is easily seen to be a Hausdorff linear topology so the operations of $(x, y) \to x + y$ and $(\lambda, x) \to \lambda x$ of $X \times X$ to X and $\mathbb{R} \times X$ to X, respectively, are continuous. This topology is *never* metrizable and *never* complete for the infinite-dimensional normed linear spaces! Usually closures are *not* determined sequentially. Nevertheless, weak sequential convergence in special spaces often holds secrets regarding the inner nature of such spaces. To be sure, if X is a normed linear space and (x_n) is a sequence of vectors in X then we say that (x_n) **converges weakly** to $x \in X$ (sometimes denoted by $x = \text{weak-}\lim_n x_n$) if for each $x^* \in X^*$,

$$\lim_n x^*(x_n) = x^*(x).$$

It is an integral part of basic functional analysis to compute the duals of special normed linear spaces and using the specific character of the spaces involved to characterize when sequences are **weakly null** (tend to zero weakly).

We now turn our attention to life inside spaces of the form $l^\infty(Q)$, where Q is a set and $l^\infty(Q)$ denotes the normed linear space of all bounded real-valued functions x defined on Q, where x's norm is given by

$$||x||_\infty = \sup\{|x(q)| : q \in Q\}.$$

Now to bring tools like the Hahn-Banach Theorem to bear on the study of $l^\infty(Q)$, we need to know something about $l^\infty(Q)^*$. Banach knew a great deal about this (as did F. Riesz before him); he didn't formulate an *exact* description of $l^\infty(Q)^*$ but nevertheless, understood the basics. To begin, if $x^* \in l^\infty(Q)^*$, then x^* is entirely determined by its values at members of $l^\infty(Q)$ of the form χ_E where $E \subseteq Q$; after all, simple functions are dense in $l^\infty(Q)$. Now if $F(E) = x^*(\chi_E)$ then F is a bounded finitely additive real-valued measure on 2^Q, the collection of all subsets of Q. If we define $|F|$ by

$$|F|(E) = \sup\{F(S) : S \subseteq E\}$$

for $E \subseteq Q$ then $|F|$ is a non-negative real-valued function defined on 2^Q and $|F(E)| \leq |F|(E)$ for each $E \subseteq Q$. (Remember: $\chi_\emptyset = 0$ so $F(\emptyset) = x^*(\chi_\emptyset) = x^*(0) = 0$.) What's more, $|F|$ is also finitely additive! If $G \subseteq E_1 \cup E_2$ where E_1 and E_2 are disjoint subsets of Q then $G = (G \cap E_1) \cup (G \cap E_2)$, and so

$$\begin{aligned} F(G) &= F((G \cap E_1) \cup (G \cap E_2)) \\ &= F(G \cap E_1) + F(F \cap E_2) \subseteq |F|(E_1) + |F|(E_2); \end{aligned}$$

It follows that

$$|F|(E_1 \cup E_2) \leq |F|(E_1) + |F|(E_2).$$

February 3, 2009

On the other hand, if E_1 and E_2 are disjoint subsets of Q then for any $\epsilon > 0$ we can pick $G_1 \subseteq E_1$ and $G_2 \subseteq E_2$ so

$$|F|(E_1) \leq F(G_1) + \frac{\epsilon}{2},\ |F|(E_2) \leq F(G_2) + \frac{\epsilon}{2}.$$

But now

$$\begin{aligned} |F|(E_1) + |F|(E_2) &\leq F(G_1) + \frac{\epsilon}{2} + F(G_2) + \frac{\epsilon}{2} \\ &= F(G_1) + F(G_2) + \epsilon \\ &= F(G_1 \cup G_2) + \epsilon \ \text{ (since } F \text{ is finitely additive)} \\ &\leq |F|(E_1 \cup E_2) + \epsilon; \end{aligned}$$

since $\epsilon > 0$ was arbitrary, we see that

$$|F|(E_1) + |F|(E_2) \leq |F|(E_1 \cup E_2)$$

whenever E_1 and E_2 are disjoint subsets of Q.

Here's the punch line: if $x^* \in l^\infty(Q)^*$ then x^* defines a bounded finitely additive measure on 2^Q - call this measure F. From F we generate $|F|$, all of whose values are non-negative; $|F| - F$ is also a non-negative bounded real-valued finitely additive map on 2^Q and $F = |F| - (|F| - F)$. So F is the difference of *non-negative* bounded finitely additive maps on 2^Q. In turn, such non-negative additive bounded maps on 2^Q define positive linear functionals on $l^\infty(Q)$, functionals that are necessarily bounded linear functionals.

Why is this last statement so? Well suppose $G : 2^Q \to [0, \infty)$ is finitely additive. If $A \subseteq B \subseteq Q$ then $B = A \cup (B \backslash A)$ so

$$G(B) = G(A \cup (B \backslash A)) = G(A) + G(B \backslash A) \geq G(A);$$

it follows that for any $E \subseteq Q$, $G(E) \leq G(Q)$, and G is bounded by $G(Q)$. Moreover if $E_1, \dots E_k$ are pairwise disjoint subsets of Q and $a_1, \dots, a_n \in \mathbb{R}$ then

$$\begin{aligned} \left| \sum_{i \leq n} a_i G(E_i) \right| &\leq \sum_{i \leq n} |a_i| G(E_i) \leq \sup_{1 \leq i \leq n} |a_i| \sum_{i \leq n} G(E_i) \\ &= \left\| \sum_{i \leq n} a_i \chi_{E_i} \right\|_\infty G\left(\bigcup_{i \leq n} E_i \right) \\ &\leq G(Q) \left\| \sum_{i \leq n} a_i \chi_{E_i} \right\|_\infty \end{aligned}$$

and so G determines a linear functional

$$\sum a_i \chi_{E_i} \to \sum a_i G(E_i)$$

on the simple functions which is 'bounded' there on. This bounded linear functional extends to $l^\infty(Q)$ in a bounded linear fashion, a positive functional to be sure.

Now $x \in B_{l^\infty(Q)}$ means $|x(q)| \leq 1$ for all $q \in Q$; it follows that $-1 \leq x(q) \leq 1$ for all $q \in Q$.

If f is a positive linear functional defined on $l^\infty(Q)$, so $f(x) \geq 0$ whenever $x(q) \geq 0$ for all $q \in Q$, then $f(-1) \leq f(x) \leq f(1)$, or $|f(x)| \leq f(1)$, whenever $x \in B_{l^\infty(Q)}$. Positive linear functionals on $l^\infty(Q)$ are *bounded* linear functionals.

THEOREM 1.0.35 (Banach). *Let Q be a (non-empty) set and (x_n) be a (uniformly) bounded sequence in $l^\infty(Q)$. Then (x_n) is weakly null if and only if*

$$\lim_n \liminf_k |x_n(q_k)| = 0 \tag{1.3}$$

for each sequence (q_k) of points in Q.

PROOF. Necessity: Suppose to the contrary that there is a sequence of points (q_k) in Q such that

$$\limsup_n \liminf_k |x_n(q_k)| > \alpha > 0$$

for some α. Then unraveling the meaning of lim sup's, we can find a strictly increasing sequence (n_j) of positive integers such that

$$\liminf_k |x_{n_j}(q_k)| > \alpha > 0$$

for each j. Now turning to the meaning of lim inf, we find a subsequence (q_{k_m}) of (q_k) such that

$$|\lim_m x_{n_j}(q_{k_m})| > \alpha > 0$$

for each j. Let $x^* \in l^\infty(Q)$ be given by

$$x^*(x) = \mathrm{LIM}((x(q_{k_m})_m))$$

where LIM $\in l^\infty(\mathbb{N})^*$ is a Banach limit. Then for each j,

$$|x^*(x_{n_j})| > \alpha$$

and so

$$\limsup_n |x^*(x_n)| > \alpha > 0.$$

It follows that (x_n) is *not* weakly null in $l^\infty(Q)$.

Sufficiency: By remarks preceding the statement of this theorem, to test (x_n)'s weak nullity it suffices to check the action of $x^* \in l^\infty(Q)^*$, for x^* a positive linear functional of norm 1, on the sequence (x_n). So suppose x^* is such a functional, with

$$\limsup_n x^*(x_n) > \alpha > 0$$

where (1.3) holds:

$$\lim_n \liminf_k |x_n(q_k)| = 0$$

for each sequence (q_k) of points in Q. Let s_n be the sequence

$$s_n(q) = \begin{cases} x_n(q) & \text{if } x_n(q) \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

and let $t_n = x_n - s_n$.

One of $\limsup_n x^*(s_n)$ and $\limsup_n x^*(t_n)$ must exceed $\frac{\alpha}{2}$; after all, $x_n = s_n + t_n$ so

$$x^*(x_n) = x^*(s_n) + x^*(t_n)$$

ensuring that

$$\limsup_n x^*(x_n) = \limsup_n (x^*(q_n) + x^*(t_n)) \le \limsup_n x^*(s_n) + \limsup_n x^*(t_n).$$

If we clip off the 'bottoms' of s_n by defining

$$y_n(q) = \begin{cases} s_n(q) & \text{if } s_n(q) \ge \frac{\alpha}{6} \\ 0 & \text{otherwise} \end{cases}$$

then

$$||s_n - y_n||_\infty \le \frac{\alpha}{6};$$

what's more,

$$\begin{aligned} \limsup_n x^*(y_n) &= \limsup_n x^*(s_n - (s_n - y_n)) \\ &= \limsup_n (x^*(s_n) - x^*(s_n - y_n)) \end{aligned}$$

so that

$$\limsup_n x^*(y_n) \ge \frac{\alpha}{2} - \frac{\alpha}{6} = \frac{\alpha}{3}.$$

Let

$$S_n = \left\{ q \in Q : |x_n(q)| \ge \frac{\alpha}{6} \right\},$$

and look at χ_{S_n}. Since

$$||y_n||_\infty \le ||s_n||_\infty \le ||x_n||_\infty \le M,$$

say, we see that for any $q \in Q$

$$\chi_{S_n}(q) \ge \frac{y_n(q)}{M}$$

so that

$$x^*(M \cdot \chi_{S_n}) \ge x^*(y_n).$$

From this we see that

$$\limsup_n x^*(\chi_{S_n}) \ge \frac{\alpha}{3M} =: \beta > 0.$$

For $E \subseteq Q$, let $F(E) = x^*(\chi_E)$; of course $F \in l^\infty(Q)^\star$ and

$$\limsup_n F(S_n) > \beta.$$

Let n_1 be the smallest positive integer such that

$$\limsup_n F(S_n \cap S_{n_1}) > 0.$$

Such an n ***exists by the way! This is the crux of the matter!*** In fact, otherwise,

$$\lim_n F(S_n \cap S_k) = 0$$

for each k so (because F is additive)

$$\lim_n F(\cup_{j=1}^k (S_n \cap S_j) = 0,$$

for each k as well. Let $k_1 = 1$. Pick $m_1 > k_1$ so large that

$$F(S_{m_1}) > \beta$$

and

$$F(S_{m_1} \cap S_{k_1}) < \frac{\beta}{2}.$$

Let $k_2 > m_1$. Pick m_2 so large that

$$F(S_{m_2}) > \beta$$

and

$$F(S_{m_2} \cap (S_1 \cup \cdots \cup S_{k_2})) < \frac{\beta}{2}.$$

Continuing in this fashion, producing $k_1 < m_1 < k_2 < m_2 < \cdots$, with

$$F(S_{m_j}) > \beta$$

and

$$F(S_{m_j} \cap (S_1 \cup \cdots \cup S_{k_j})) < \frac{\beta}{2}.$$

Now disjointivity: let T_j be given by

$$T_j = S_{m_j} \backslash [S_{m_j} \cap (\cup_{i=1}^{k_j} S_i)].$$

By construction

$$F(T_j) > \frac{\beta}{2}$$

and this is a "no-no" since F takes disjoint sequences to 0.

So n_1 does indeed exist such that

$$\limsup_n F(S_n \cap S_{n_1}) > 0.$$

Believe it or not.

Once faith has been established for n_1 we see that there are $n_2 < n_3 < \cdots < n_k < \cdots$ so that

$$\limsup_n F(S_n \cap S_{n_1} \cap \cdots \cap S_{n_k}) > 0$$

for each k. The all-important point here is that for each k there is at least one point q_k so

$$q_k \in S_{n_1} \cap S_{n_2} \cap \cdots \cap S_{n_k}.$$

Of course if $j \geq k$ then $q_j \in S_{n_k}$ and so by how the S_n's were defined

$$|x_{n_k}(q_j)| \geq \frac{\alpha}{6}.$$

It soon follows that

$$\limsup_n \liminf_j |x_n(q_j)| \geq \frac{\alpha}{6}.$$

This contradicts (1.3) and thus, the sufficiency is proven. □

This result is remarkable because it characterizes weak convergence in a highly non-separable, non-metrizable situation in terms of *sequences* q_k in Q.

Suppose Q is a non-void compact metric space. Then the space $C(Q)$ of all continuous functions real-valued functions defined on Q, equipped with the norm

$$||x||_\infty = \sup\{|x(q)| : q \in Q\}$$

is a closed linear subspace of $l^\infty(Q)$, the space of all bounded real-valued functions on Q. We can give a much more succinct characterization of when a bounded sequence (x_n) is weakly null in $C(Q)$ then just what Banach has above. Indeed, and again this was observed by Banach.

THEOREM 1.0.36 (Banach). *A (uniformly) bounded sequence (x_n) in $C(Q)$ is weakly null if and only if*

$$\lim_n x_n(q) = 0$$

for each $q \in Q$.

PROOF. If $(x_n) \subseteq C(Q)$ is weakly null and $q \in Q$ then using the point evaluation

$$\delta_q \in C(K)^*, \text{where } \delta_q(x) = x(q)$$

we see that

$$0 = \lim_n \delta_q(x_n) = \lim_n x_n(q).$$

Now assume that (x_n) is a (uniformly) bounded sequence in $C(Q)$ for which $\lim_n x_n(q) = 0$ for each $q \in Q$ and imagine that (x_n) is *not* weakly null in $C(Q)$. Of course the Hahn-Banach theorem tells us that (x_n) is not weakly null in $l^\infty(Q)$ either. So by Theorem 1.0.35 there must be a subsequence $(x_n^{'})$ of (x_n) and a sequence of points (q_k) in Q and an $\alpha > 0$ so that for each n

$$\liminf_k |x_n^{'}(q_k)| \geq \alpha > 0. \tag{1.4}$$

But (q_k) is a sequence in the compact (hence sequentially compact) metric space Q and so (q_n) has a subsequence $(q_k^{'})$ that converges to some $q_0 \in Q$. By (1.4) it must be that for all n

$$|x_n^{'}(q_0)| \geq \alpha > 0.$$

OOPS. □

2. The Lebesgue Integral on Abstract Spaces

In this section we present Banach's approach to the Lebesgue integral in abstract spaces. Banach's approach is a Daniell-like construction built using his clear and deep understanding of lim sup's and lim inf's. His starting point is a positive linear functional f acting on a vector lattice $\mathcal{C}$ of real-valued functions defined on some set K; we suppose (with Banach) that the functional satisfies a kind of Bounded Convergence Theorem (BCθ) on $\mathcal{C}$. It is important to note that in the previous section we presented Banach's famous result characterizing weakly convergent sequences in spaces $C(Q)$, Q a compact metric space; it follows from this that should the initial vector lattice $\mathcal{C}$ be such a $C(Q)$, then every positive linear functional satisfies the (BCθ) hypothesis.

After an initial discussion of technical consequences of the (BCθ) hypothesis involving lim sup's and lim inf's of functions in $\mathcal{C}$, Banach introduces an upper and a lower integral. Were we doing measure theory, this piece of the puzzle would be concerned with properties of outer and inner measures generated from an initial set function.

Next the class of integrable functions is isolated, being identified as those real-valued functions for which the upper and lower integrals coincide and are simultaneously finite. The classical Monotone and Dominated Convergence Theorems are derived and all is well with the world.

We follow a discussion of what the construction does in the all important case that the initial vector lattice $\mathcal{C} = C(Q)$, Q a compact metric space.

It is noteworthy that this construction of Banach led him to a description of $C(Q)^*$. It is

February 3, 2009

impossible to tell for certain (but easy to imagine) what Saks thought of Banach's 'construction' of $C(Q)^*$. It appeared after all, as an appendix to Saks' classic monograph "*Theory of the Integral*", yet uses **practically *none*** of the material from the monograph! Whatever Saks though, he soon presented an elegant proof of Banach's result about $C(Q)^*$ in the Duke Mathematics Journal. We present Saks' proof in an appendix to this chapter.

2.1. A Start. The numbering and notation throughout this section is consistent with Banach's. Let $\mathcal{C}$ denote a vector lattice of real-valued functions defined on a set Q. (i.e., If $x, y \in \mathcal{C}$ then so is $x \vee y = \inf\{x, y\}$ and $x \wedge y = \sup\{x, y\}$.) A linear functional f on $\mathcal{C}$ is a *positive linear functional* if $f(x) \geq 0$ for any $x \in \mathcal{C}, x \geq 0$.

Throughout this section, we will suppose f is a positive linear functional on $\mathcal{C}$ satisfying

$$\text{BC}\boldsymbol{\theta} \begin{cases} \text{if } \boldsymbol{(x_n) \subseteq \mathcal{C}, M \in \mathcal{C},} \text{ with } \boldsymbol{|x_n| \leq M} \text{ and } \boldsymbol{\lim_n x_n(t) = 0} \text{ for all } \boldsymbol{t \in Q,} \text{ then} \\ \boldsymbol{\lim_n f(x_n) = 0.} \end{cases}$$

If $\mathcal{C} = C(Q)$, Q a compact metric space then **any** positive linear functional f on $\mathcal{C}$ satisfies (BCθ) thanks to Theorem 1.0.36. This is an example well worth keeping in mind.

1° *If* $(x_n) \subseteq \mathcal{C}, m \in \mathcal{C}, z \geq 0, x_n \geq m,$ *and* $\liminf_n x_n \geq z$ *then*

$$\lim_n x_n - |x_n| = 0.$$

Now $\liminf_n x_n \geq z$ ensures $\liminf_n x_n(t) \geq 0$ for each $t \in Q$, that is, for each $t \in Q$,

$$\lim_{n \to \infty} \inf\{x_n(t), x_{n+1}(t), \ldots\} \geq 0.$$

So, given $\eta > 0$ there's $n = n(\eta)$ so that for all $k \geq n$,

$$x_k(t) \geq -\eta. \tag{2.1}$$

Look at $x_n - |x_n|$:

$$(x_n - |x_n|)(t) = \begin{cases} 0, & \text{if } x_n(t) \geq 0 \\ 2x_n(t), & \text{if } x_n(t) < 0; \end{cases}$$

naturally, $(x_n - |x_n|)(t) \leq 0$ for all t. So, if $\lim_n (x_n - |x_n|)$ exists, it must be less than or equal to 0.

Let's check on contrary possibilities. Can it be that

$$\liminf_n (x_n - |x_n|)(t_0) < 0$$

for some $t_0 \in Q$? Let's suppose that this is possible. Notice since $m \in \mathcal{C}$

$$\liminf_n (x_n - |x_n|)(t_0) \geq \inf\{2m(t_0), 0\} > -\infty.$$

If

$$\liminf_n (x_n - |x_n|)(t_0) < 0$$

it's because there's a subsequence $(x_n^{'})$ of x_n so that for some $\epsilon_0 > 0$

$$2x_n^{'}(t_0) = x_n^{'}(t_0) - |x_n^{'}(t_0)| < -\epsilon_0,$$

for all n. It follows that

$$x_n^{'}(t_0) < \frac{-\epsilon_0}{2}$$

for all n. But if n is BIG we can arrange

$$x_n'(t_0) > \frac{-\epsilon_0}{4}$$

(that's what our opening words (2.1) of wisdom guarantee). Ah ha! While

$$x_n'(t_0) > \frac{-\epsilon_0}{4},$$

we also have

$$x_n'(t_0) < \frac{-\epsilon_0}{2}.$$

Drawing the conclusion that $\frac{\epsilon_0}{2} < \frac{\epsilon_0}{4}$ is easy from this and leaves us with a clear-cut contradiction to all that is right in our world. We conclude that

$$\liminf_n (x_n - |x_n|) \geq 0.$$

We know that

$$\limsup_n (x_n - |x_n|) \leq 0,$$

and so 1° follows. □

2° *If* $(x_n) \subseteq \mathcal{C}, m \in \mathcal{C}, x_n \geq m,$ *and* $\liminf_n x_n \geq 0$ *then*

$$\liminf_n f(x_n) \geq 0.$$

Again start with a look at $x_n - |x_n|$; as before since $2m \in \mathcal{C}$

$$-\infty < \inf\{2m, 0\} \leq x_n - |x_n| \leq 0.$$

By 1°, we know that

$$\lim_n (x_n - |x_n|)(t) = 0$$

for each $t \in Q$, and since $\mathcal{C}$ is a lattice, we can use the the (BCθ) condition on f to conclude

$$\lim_n f(x_n - |x_n|) = 0.$$

Can $\liminf_n f(x_n) < 0$? If so then it's because there is a subsequence (x_n') of (x_n) and an $\epsilon_0 > 0$ so

$$f(x_n') < -\epsilon_0$$

for all n. We can (and do) assume that

$$\lim_n f(x_n')$$

exists as well. But now

$$\lim_n f(x_n'), \lim_n f(x_n' - |x_n'|)$$

both exist and so

$$\lim_n f(|x_n'|)$$

exists, too, with

$$\begin{aligned} \lim_n f(|x_n'|) &= \lim_n (f(|x_n'| - x_n') + f(x_n')) \\ &= \lim_n f(x_n') \leq -\epsilon_0, \end{aligned}$$

which is not possible, and so we have 2°. □

2.2. Upper and lower integrals. Let $\mathcal{L}^*$ denote the set of all real-valued functions z on Q for which there exist two sequences $(x_n), (y_n) \subseteq \mathcal{C}$ such that

$$\limsup_n y_n \leq z \leq \liminf_n x_n.$$

$\mathcal{L}^*$ is a linear space containing $\mathcal{C}$ since $\mathcal{C}$ is a vector lattice.

Given $z \in \mathcal{L}^*$, the *upper integral of* z, $\overline{\int}(z)$, is defined by

$$\overline{\int}(z) = \inf\{\liminf_n f(x_n) : \text{there exists } m \in \mathcal{C}, (x_n) \subseteq \mathcal{C}, x_n \geq m, \liminf_n x_n \geq z\};$$

the *lower integral of* z, $\underline{\int}(z)$, is defined by

$$\underline{\int}(z) = \sup\{\limsup_n f(x_n) : \text{there exists } M \in \mathcal{C}, (x_n) \subseteq \mathcal{C}, x_n \leq M, \limsup_n x_n \leq z\}.$$

Obviously

$$\underline{\int} z = -\overline{\int}(-z).$$

Note: in each of the above definitions we can suppose that $\lim_n f(x_n)$ exists and is real valued since $\liminf$'s and $\limsup$'s are taken over sequences which are eventually finite. So we replace $\liminf_n f(x_n)$ and $\limsup_n f(x_n)$ with $\lim_n f(x_n)$ throughout.

From the definitions we have

$3°$ *If* $z \in \mathcal{L}^*, z \geq 0$, *and* $\bar{\int}(z) < P < \infty$ *then we can find* $(x_n) \subseteq \mathcal{C}, x_n \geq 0, \liminf x_n \geq z$ *with* $f(x_n) < P$ *for all* n.

The value of $3°$ is found in the accessibility it affords us to epsilonics; since (with Banach) we have frequent call on computing $\limsup$'s and $\liminf$'s. This is a critical aid.

LEMMA 2.2.1. *For any* $x \in \mathcal{C}, \bar{\int}(x) = f(x)$.

PROOF. On the one hand, we can let $x_n = x$ for all n and $m = x$; this done, we plainly have

$$\liminf_n x_n \geq x, \quad \text{and} \quad x_n \geq m.$$

Hence

$$\overline{\int}(x) \leq \liminf_n f(x_n) = f(x).$$

On the other hand, if $(x_n) \subseteq \mathcal{C}$ and $m \in \mathcal{C}$ with

$$\liminf_n x_n \geq x, \quad \text{and} \quad x_n \geq m,$$

then

$$\liminf_n (x_n - x) \geq 0, \quad \text{and} \quad x_n - x \geq m - x.$$

$2°$ steps in to say

$$\text{“}0 \leq \liminf_n f(x_n - x) = \liminf_n f(x_n) - f(x)\text{”};$$

we see that

$$f(x) \leq \liminf_n f(x_n)$$

and with this we conclude

$$f(x) \le \overline{\int}(x).$$

$\square$

LEMMA 2.2.2. *If $z_1, z_2 \in \mathcal{L}^*$ with $\overline{\int}(z_1), \overline{\int}(z_2) < \infty$, then $\overline{\int}(z_1 + z_2) \le \overline{\int}(z_1) + \overline{\int}(z_2)$.*

PROOF. Suppose P_1, P_2 are numbers such that

$$\overline{\int}(z_1) < P_1, \text{ and } \overline{\int}(z_2) < P_2.$$

There are sequences $(x_n^{(1)}), (x_n^{(2)}) \subseteq \mathcal{C}$ and functions $m_1, m_2 \in \mathcal{C}$ such that

$$\begin{aligned} \liminf_n x_n^{(1)} &\ge z_1, \qquad x_n^{(1)} \ge m_1 \text{ for all } n, \\ \liminf_n x_n^{(2)} &\ge z_2, \qquad x_n^{(2)} \ge m_2 \text{ for all } n, \end{aligned}$$

and

$$\lim_n f(x_n^{(1)}) < P_1, \ \lim_n f(x_n^{(2)}) < P_2.$$

Letting $x_n = x_n^{(1)} + x_n^{(2)}$ and $m = m_1 + m_2$, we see that

$$\liminf_n x_n \ge z_1 + z_2, \ x_n \ge m.$$

It follows that

$$\overline{\int}(z_1 + z_2) \le \lim_n f(x_n) = \lim_n f(x_n^{(1)}) + \lim_n f(x_n^{(2)}) < P_1 + P_2.$$

Enough said. $\square$

LEMMA 2.2.3. *For any $z \in \mathcal{L}^*$,*

$$\underline{\int}(z) \le \overline{\int}(z).$$

PROOF. There is nothing to prove if $\overline{\int}(z) = +\infty$. If $\overline{\int}(-z) = +\infty$ then $\underline{\int}(z) = -\overline{\int}(-z)$ so again there is nothing to prove. If $\overline{\int}(z), \overline{\int}(-z) < \infty$ then Lemma 2.2.2 kicks in to give

$$0 = f(0) = \overline{\int}(0) = \overline{\int}(z - z) \le \overline{\int}(z) + \overline{\int}(-z)$$

so that

$$\underline{\int}(z) = -\overline{\int}(-z) \le \overline{\int}(z).$$

$\square$

LEMMA 2.2.4. *If $z \in \mathcal{L}^*$ and $\overline{\int}(z) < \infty$, then*

$$\overline{\int}\left(\frac{z + |z|}{2}\right) < \infty,$$

and

$$\overline{\int}(z) = \overline{\int}\left(\frac{z + |z|}{2}\right) + \overline{\int}\left(\frac{z - |z|}{2}\right).$$

PROOF. Suppose $\bar{\int}(z) < P < \infty$. Find $m \in \mathcal{C}$ and $(x_n) \subseteq \mathcal{C}$ so that $x_n \geq m$ for all n, and $\liminf_n x_n \geq z$, $\lim_n f(x_n) < P$. Notice that if $x_n \geq m$ then

$$\frac{x_n - |x_n|}{2} \geq \frac{m - |m|}{2};$$

this can be seen by a simple analysis of cases.

Hence

$$\begin{aligned} f\left(\frac{x_n + |x_n|}{2}\right) &= f\left(x_n - \left(\frac{x_n - |x_n|}{2}\right)\right) \\ &= f(x_n) - f\left(\frac{x_n - |x_n|}{2}\right) \\ &\leq f(x_n) - f\left(\frac{m - |m|}{2}\right) \end{aligned}$$

and

$$\begin{aligned} \bar{\int}\left(\frac{z + |z|}{2}\right) &\leq \liminf_n f\left(\frac{x_n + |x_n|}{2}\right) \\ &\leq \lim_n f(x_n) - f\left(\frac{m + |m|}{2}\right) \\ &< \infty. \end{aligned}$$

Now

$$\begin{aligned} P &> \lim_n f(x_n) \\ &\geq \liminf_n f\left(\frac{x_n + |x_n|}{2}\right) + \liminf_n f\left(\frac{x_n - |x_n|}{2}\right) \\ &\geq \bar{\int}\left(\frac{z + |z|}{2}\right) + \bar{\int}\left(\frac{z - |z|}{2}\right); \end{aligned}$$

it follows from P's arbitrary nature among members $> \bar{\int}(z)$ that

$$\bar{\int}(z) \geq \bar{\int}\left(\frac{z + |z|}{2}\right) + \bar{\int}\left(\frac{z - |z|}{2}\right).$$

Lemma 2.2.2 tells the rest of this tale. □

Two more lemmas are plain and worth mentioning.

LEMMA 2.2.5. *If* $z_1, z_2 \in \mathcal{L}^*$ *satisfy* $z_1 < z_2$ *then* $\bar{\int}(z_1) \leq \bar{\int}(z_2)$*; in particular if* $z \in \mathcal{L}^*$ *and* $z \geq 0$ *then* $\bar{\int}(z) \geq 0$.

LEMMA 2.2.6. *If* $z \in \mathcal{L}^*$ *then* $\bar{\int}(\lambda z) = \lambda \bar{\int}(z)$ *for any real number* $\lambda \geq 0$.

2.3. The Integral. Let $\mathcal{L}$ be the set

$$\mathcal{L} = \{z \in \mathcal{L}^* : \bar{\int}(z) = \underline{\int}(z), \text{ with both finite}\}.$$

LEMMA 2.3.1. $\mathcal{L}$ *is a linear space and* $\bar{\int}$ *is a linear functional on* $\mathcal{L}$*. Moreover* $\mathcal{C} \subseteq \mathcal{L}$ *and* $\bar{\int}$ *extends* f.

Lemma 2.3.2. *If $z \in \mathcal{L}$ then $|z| \in \mathcal{L}$, that is, $\mathcal{L}$ is a vector lattice.*

Proof. Since

$$|z| = \left(\frac{z+|z|}{2}\right) + \left(\frac{|z|-z}{2}\right),$$

it's enough (thanks to $\mathcal{L}$'s linearity) to show that $\frac{z+|z|}{2}, \frac{z-|z|}{2}$ both belong to $\mathcal{L}$ if $z \in \mathcal{L}$. We recall that Lemma 2.2.4 ensures that if $z \in \mathcal{L}$ then $\overline{\int}\left(\frac{z+|z|}{2}\right) < \infty$ and $\overline{\int}\left(\frac{z-|z|}{2}\right) > -\infty$, as well as

$$\overline{\int}(z) = \overline{\int}\left(\frac{z+|z|}{2}\right) + \overline{\int}\left(\frac{z-|z|}{2}\right).$$

Symmetry (applying Lemma 2.2.4 to $-z$ and using $\overline{\int} - z = \underline{\int} z$) shows that $\underline{\int}\left(\frac{z+|z|}{2}\right) < \infty$ and $\underline{\int}\left(\frac{z-|z|}{2}\right) > -\infty$ as well as

$$\underline{\int}(z) = \underline{\int}\left(\frac{z+|z|}{2}\right) + \underline{\int}\left(\frac{z-|z|}{2}\right).$$

But $z \in \mathcal{L}$ says that $\overline{\int} z = \underline{\int} z$ and so

$$\begin{aligned}\left[\overline{\int}\left(\frac{z+|z|}{2}\right) - \underline{\int}\left(\frac{z+|z|}{2}\right)\right] + \left[\overline{\int}\left(\frac{z-|z|}{2}\right) - \underline{\int}\left(\frac{z-|z|}{2}\right)\right] &= \\ \left[\overline{\int}\left(\frac{z+|z|}{2}\right) + \overline{\int}\left(\frac{z-|z|}{2}\right)\right] - \left[\underline{\int}\left(\frac{z+|z|}{2}\right) + \underline{\int}\left(\frac{z-|z|}{2}\right)\right] &= \overline{\int} z - \underline{\int} z = 0.\end{aligned}$$

Now Lemma 2.2.3 kicks in to say that the *finite* quantities $\overline{\int}\left(\frac{z+|z|}{2}\right), \underline{\int}\left(\frac{z+|z|}{2}\right)$ must, in fact, be equal and the finite quantities $\overline{\int}\left(\frac{z-|z|}{2}\right), \underline{\int}\left(\frac{z-|z|}{2}\right)$ follow suit. □

An old friend is next on the agenda.

Lemma 2.3.3 (Monotone Convergence Theorem).

$$MC\theta \begin{cases} \textit{if } (z_n) \subseteq \mathcal{L}, z_n \le z_{n+1} \textit{ for all } n \textit{ and } z = \lim_n z_n \textit{ with} \\ \textit{lim}_n \overline{\int}(z_n) < \infty, \ \textit{then } z \in \mathcal{L} \textit{ and } \overline{\int}(z) = \lim_n \overline{\int}(z_n). \end{cases}$$

Proof. We can, and do, assume $z_1 = 0$; otherwise subtract z_1 from each of the z_n's. Next note that $z \ge z_n$ for all n and so since $\overline{\int}(z_n) = \underline{\int}(z_n)$ we have

$$\underline{\int}(z) \ge \lim_n \underline{\int}(z_n) = \lim_n \overline{\int}(z_n). \tag{2.2}$$

Let $\epsilon > 0$. Let $w_n = z_{n+1} - z_n \ge 0$. For each n find $(w_k^{(n)}) \subseteq \mathcal{C}$ with $w_k^{(n)} \ge 0, \lim_k w_k^{(n)} \ge w_n$ and

$$f(w_k^{(n)}) \le \overline{\int}(w_n) + \frac{\epsilon}{2^n},$$

say. Write

$$y_n = w_n^{(1))} + \cdots + w_n^{(n)};$$

so

$$\begin{aligned} y_1 &= w_1^{(1)} \in \mathcal{C} \\ y_2 &= w_2^{(1)} + w_2^{(2)} \in \mathcal{C} \\ &\vdots \\ y_n &= w_n^{(1)} + w_n^{(2)} + \cdots + w_n^{(n)} \in \mathcal{C} \end{aligned}$$

Since $y_n = w_n^{(1)} + w_n^{(2)} + \cdots + w_n^{(n)} \geq w_1 + w_2 + \cdots w_n$,

$$\liminf_n y_n \geq \liminf_n (w_1 + w_2 + \cdots + w_n) = \sum_n w_n = z.$$

At the same time,

$$\begin{aligned} f(y_n) &= f(w_n^{(1)}) + \cdots + f(w_n^{(n)}) \\ &\leq \overline{\int}(w_1) + \frac{\epsilon}{2} + \overline{\int}(w_2) + \frac{\epsilon}{2^2} + \cdots + \overline{\int}(w_n) + \frac{\epsilon}{2^n} \\ &< \overline{\int}(w_1) + \cdots + \overline{\int}(w_n) + \epsilon \\ &= \overline{\int}(z_2) + \overline{\int}(z_3 - z_2) + \cdots + \overline{\int}(z_{n+1} - z_n) + \epsilon \\ &= \overline{\int}(z_2 + (z_3 - z_2) + \cdots + (z_{n+1} - z_n)) + \epsilon \\ &= \overline{\int}(z_{n+1}) + \epsilon \leq \lim_n \overline{\int}(z_n) + \epsilon. \text{ (by Lemma 2.2.5)} \end{aligned}$$

It follows that

$$\overline{\int}(z) \leq \liminf_n f(y_n) \leq \lim_n \overline{\int}(z_n) + \epsilon,$$

and by ϵ's arbitrary nature,

$$\begin{aligned} 0 &\leq \overline{\int}(z) \leq \lim_n \overline{\int}(z_n) \\ &\leq \underline{\int}(z) \text{ by (2.2)} \\ &\leq \overline{\int}(z) < \infty. \end{aligned}$$

Therefore

$$\overline{\int}(z) = \lim_n \overline{\int}(z_n).$$

□

Another old friend.

LEMMA 2.3.4 (Dominated Convergence Theorem). *Suppose* $(z_n) \subseteq \mathcal{L}, M \in \mathcal{L}$ *and* $|z_n| \leq M$. *Then*

$$g = \liminf_n z_n, \; h = \limsup_n z_n \in \mathcal{L}$$

with

$$\overline{\int}(g) \leq \liminf_n \overline{\int}(z_n) \leq \limsup_n \overline{\int}(z_n) \leq \overline{\int}(h).$$

Consequently

$$DC\theta \begin{cases} \textit{Suppose } (z_n) \subseteq \mathcal{L}, M \in \mathcal{L} \textit{ satisfy } |z_n| \leq M. \\ \textit{If } z(t) = \lim_n z_n(t) \textit{ for each } t \in Q \textit{ then } z \in \mathcal{L} \textit{ and } \overline{\int}(z) = \lim_n \overline{\int}(z_n). \end{cases}$$

Proof. For each i and for each $j \geq i$, write

$$g_{ij} = \min\{z_i, z_{i+1}, \ldots, z_j\}.$$

Then the sequence $(g_{ij})_{j=i}^{\infty}$ is decreasing, each member belongs to $\mathcal{L}$ and so the sequence $(M - g_{ij})_{j=i}^{\infty}$ is an increasing sequence of members of $\mathcal{L}$. (MCθ) guarantees that if

$$g_i = \lim_j g_{ij},$$

then

$$M - g_i \in \mathcal{L}$$

and

$$\overline{\int}(M - g_i) = \lim_j \overline{\int}(M - g_{ij});$$

that is, $g_i \in \mathcal{L}$ and

$$\overline{\int}(g_i) = \lim_j \overline{\int}(g_{ij}).$$

Applying (MCθ) again reveals

$$g = \liminf_n z_n = \lim_j g_{ij} \in \mathcal{L}$$

with

$$\overline{\int}(g) = \lim_i \overline{\int}(g_i) \leq \liminf_n \overline{\int}(z_n).$$

□

Lemma 2.3.5. *If $z \in \mathcal{L}, z \geq 0$ and $\overline{\int} z = 0$ then whenever the function x satisfies $|x| \leq z$ we have that $x \in \mathcal{L}$ and $\overline{\int} x = 0$.*

This is an immediate consequence of lemma 2.2.5.

2.4. We have the integral... Now let's turn to Banach's approach to integration. Start with $\mathcal{C} = C(Q)$, Q a compact metric space. If f is a positive linear functional on $\mathcal{C}$ then $|f(x)|$ is bounded by $f(1)$ so long as $x \in B_{\mathcal{C}}$. Why? Well if $x \in B_{\mathcal{C}}$ then $|x| \leq 1$ and $|x(q)| \leq 1$ for all $q \in Q$. So $x \in B_{\mathcal{C}}$ means $-1 \leq x(q) \leq 1$; from this it follows that $-f(1) = f(-1) \leq f(x) \leq f(1)$. Okay? *Any positive linear functional f on $C(Q)$ is a bounded linear functional with norm $f(1)$.* By Banach's theorem (Theorem 1.0.36), if $(x_n) \subseteq B_{\mathcal{C}}$ and $|x_n| \leq M \in \mathcal{C}$ with $\lim_n x_n(q) = 0$ for all $q \in Q$, then (x_n) is weakly null in $C(Q)$, hence $\lim_n f(x_n) = 0$. This is (BCθ) in Banach's integration theory!

So we can extend f to a vector lattice $\mathcal{L}$ of real-valued functions defined on Q in a linear fashion so that the extension, denoted by $\int df$, enjoys the fruits of (MCθ) and (DCθ), that is

> if $(z_n) \subseteq \mathcal{L}, (z_n)$ an ascending sequence with $z = \lim_n z_n, z \in \mathcal{L}$ whenever $\lim_n \int z_n df < \infty$; in this case, $\int z = \lim_n \int z_n$. (MCθ)
>
> Suppose $(z_n) \subseteq \mathcal{L}$ and $M \in \mathcal{L}$ satisfy $|z_n| \leq M$. Then $\liminf z_n, \limsup z_n \in \mathcal{L}$ and $\int \liminf z_n \leq \liminf \int z_n \leq \limsup \int z_n \leq \int \limsup z_n$. (DC$\theta$)

So should $(z_n) \subseteq \mathcal{L}, \mathcal{M} \in \mathcal{L}$ satisfy $|z_n| \leq \mathcal{M}$, if $z(q) = \lim_n z_n(q)$ for each $q \in Q$ then $z \in \mathcal{L}$ and $\int z = \lim_n \int z_n$.

What does this provide us with? To start, let F be a closed subset of Q. Consider the continuous function $\phi_F(q) = d(q; F)$; notice that $q \in F$ precisely when $\phi_F(q) = 0$. If we consider $\phi_n \in C(Q)$ to be

$$\phi_n(q) = \inf \left\{ \phi_F(q), \frac{1}{n} \right\},$$

then $\lim_n \phi_n(q) = \chi_F(q)$ for each $q \in Q$. Moreover, $|\phi_n| \leq 1 \in \mathcal{C} \subseteq \mathcal{L}$; hence $\chi_F \in \mathcal{L}$ and $\int \chi_F df = \lim_n \int \phi_n df$ be either (MCθ) or (DCθ), take your pick.

Suppose $E \subseteq Q$ and $\chi_E \in \mathcal{L}$. Then $1 = \chi_E + \chi_{E^c}$ so $\chi_{E^c} \subseteq \mathcal{L}$, too – afterall $1 \in \mathcal{C}$ and $\chi_E \in \mathcal{L}$. If $A, B \subseteq Q$ and $\chi_A, \chi_B \in \mathcal{L}$ then since $\chi_{A\cap B} = \inf\{\chi_A, \chi_B\}$, we see that $\chi_{A\cap B} \in \mathcal{L}$. Because

$$\chi_A + \chi_B = \chi_{A\cup B} - \chi_{A\cap B}$$

we see that should $\chi_A, \chi_B \in \mathcal{L}$ then $\chi_{A\cup B} \in \mathcal{L}$ as well. Therefore

$$\{E \subseteq Q : \chi_E \in \mathcal{L}\}$$

is an algebra of sets containing all the closed subsets of Q.

Suppose (E_n) is an ascending sequence of subsets of Q such that $\chi_{E_n} \in \mathcal{L}$ for each n. Then

$$\chi_{\cup_n E_n} = \lim_n \chi_{E_n}$$

and

$$|\chi_{E_n}| \leq 1 \in \mathcal{C} \subseteq \mathcal{L}$$

so $\chi_{\cup_n E_n} \in \mathcal{L}$ too - here you can appeal to (MCθ) or (DCθ), take your choice.

Therefore $\{E \in Q : \chi_E \in \mathcal{L}\}$ is a $\sigma-$algebra of subsets of Q containing each and every closed subset of Q. Therefore for any Borel set $B \subseteq Q, \chi_B \in \mathcal{L}$. Thus we have proved

Theorem 2.4.1. *$\mathcal{L}$ contains all of the indicator functions on Borel sets.*

Suppose (B_n) is a sequence of pairwise disjoint Borel subsets of Q. Then $\chi_{B_n}, \chi_{\cup_n B_n} \in \mathcal{L}$ and

$$\chi_{\cup_n B_n} = \sum_n \chi_{B_n};$$

an appeal to (MCθ) or (DCθ) soon reveals that

$$\int \chi_{\cup_n B_n} df = \sum_n \int \chi_{\phi_n} df.$$

$\int df$ acts in a countably additive fashion on $\mathcal{B}o(Q)$! It is a measure on $\mathcal{B}o(Q)$. i.e.,

$$\mu_f = \int df, \text{ where } \mu_f(E) = \int \chi_E df, \text{ whenever } \chi_E \in \mathcal{L}$$

is a countably additive measure defined on some $\sigma-$ field that contains the Borel $\sigma-$ field. The total mass of μ_f is $f(1)$.

3. A Brief Intermission

While Banach did not base his derivation of $C(Q)^*$ on the material of Saks' monograph, he did find it useful in his development of Haar measure for compact metric groups. In particular, he called on metric outer measures as a guiding light for his passage to Haar measure.

What follows is a presentation of the basics of metric outer measures, enough to see us through Banach's proof of the existence of an invariant (outer) measure and, afterwards, Saks' elegant proof that $C(Q)^*$ is what it is.

Preceding via the standard path, we suppose that we have a metric space Ω with metric d and a premeasure τ defined on the family $\mathcal{F}$ of subsets of Ω. Let $\delta > 0$ and denote by

$$\mathcal{F}_\delta = \{C \in \mathcal{F} : \text{ diameter } C \leq \delta\}.$$

Denote by τ_δ the restriction of τ to C_δ. The result is a premeasure that generates an outer measure μ_δ on Ω by Method I, namely for $E \subseteq \Omega$,

$$\mu_\delta := \inf\left\{\sum_i \tau_\delta(E_n) : E_n \in \mathcal{F}_\delta, E \subseteq \cup_n E_n\right\} = \inf\left\{\sum_n \tau(E_n) : E_n \in \mathcal{F}, E \subseteq \cup_n E_n, \text{diam}(E_n) \leq \delta\right\}.$$

It's plain that as δ gets smaller there are fewer members of $\mathcal{F}$ having diameter $\leq \delta$ so $\mu_\delta(E)$ gets bigger. Hence

$$\mu(E) := \sup_{\delta>0} \mu_\delta(E) = \lim_{\delta \searrow 0} \mu_\delta(E),$$

exists and is well-defined.

THEOREM 3.0.2. *μ is a measure.*

The only possible stumbling point to this is the countable subadditivity so let's see why μ is countably subadditive. To this end, let (E_n) be a sequence of subsets of Ω and consider the quantities

$$\mu(\cup_n E_n) \text{ and } \sum_n \mu(E_n).$$

Obviously the latter exceeds the former if it's $\sum_n \mu(E_n) = \infty$ so we may as well assume $\sum_n \mu(E_n) < \infty$.

Now for each $\delta > 0$, μ_δ is a known outer measure so

$$\mu_\delta(\cup_n E_n) \leq \sum_n \mu_\delta(E_n),$$

which in turn is

$$\leq \sum_n \mu(E_n).$$

February 3, 2009

The countable subadditivity follows from this. □

A key ingredient to our mix is provided in the following.

THEOREM 3.0.3. *If A and B are non-empty subsets of Ω that are 'positively separated' then*

$$\mu(A \cup B) = \mu(A) + \mu(B).$$

Here A and B 'positively separated' means that there is a $\delta > 0$ so that for any $a \in A$ and $b \in B$,

$$\rho(a, b) \geq \delta.$$

PROOF. We need to show that if A and B are positively separated then

$$\mu(A \cup B) \geq \mu(A) + \mu(B)$$

where all the terms involved are finite. The idea of the proof is to cover A, B and $A \cup B$ with very fine covers from the domain $\mathcal{F}$ of τ, a cover so fine that we can distinguish which members of the cover touch A from those that touch B.

More precisely, suppose $\epsilon_0 > 0$ is so small that

$$d(a, b) \geq \epsilon_0$$

for any $a \in A$ and $b \in B$. Let $\epsilon > 0$ announce its presence, $\epsilon < \epsilon_0/3$. Let $\epsilon^{'}, \epsilon^{''} > 0$ be such that $\epsilon^{'}, \epsilon^{''} < \epsilon$. Let $\eta = \min\{\epsilon^{'}, \epsilon^{''}\}$. Since

$$\mu(A \cup B) = \sup_{\delta > 0}\{\inf \sum_n \tau(C_n)\},$$

where the infimum is taken over all sequences (C_n) of members of $\mathcal{F}$ such that each C_n has diameter $\leq \delta$ and $A \cup B \subseteq \cup_n c_n$, we can choose (C_n) from $\mathcal{F}$ in such a way that

$$A \cup B \subseteq \cup_n C_n,$$

$$\text{diam } C_n \leq \eta,$$

and

$$\sum_n \tau(C_n) \leq \mu(A \cup B) + \epsilon.$$

By choosing C_n's this way we see that a given C_n can intersect A or B but *not both.*

Here's the first punchline; **if each C_i has diameter $\leq \eta$ then each has diameter $\leq \frac{\delta}{3}$ and so a given C_i can intersect A or B but not both.** It follows that

$$\sum_{C_i \cap A \neq \emptyset} \tau(C_i) + \sum_{C_i \cap B \neq \emptyset} \tau(B_i) \leq \sum_i \tau(C_i) \leq \mu(A \cup B) + \epsilon.$$

But each C_i that intersects A has diameter $\leq \eta \leq \delta_1$ so knowing, as we do, that A is a subset of

$$\bigcup_{C_i \cap A \neq \emptyset} C_i,$$

we get

$$\mu_{\delta_1}(A) \leq \sum_{C_i \cap A \neq \emptyset} \tau(C_i).$$

Similarly,

$$\mu_{\delta_2}(B) \leq \sum_{C_i \cap B \neq \emptyset} \tau(C_i).$$

The result:

$$\mu_{\delta_1}(A) + \mu_{\delta_2}(B) \leq \sum_{C_i \cap A \neq \emptyset} \tau(C_i) + \sum_{C_i \cap B \neq \emptyset} \tau(C_i) \leq \mu(A \cup B) + \epsilon,$$

and so Method II leads us to believe

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

Since $\epsilon > 0$ was arbitrary, this proof is done. □

The property of an outer measure on a metric space that we have isolated above is important enough to earn a special designation: an outer measure μ defined on the subsets of a metric space (Ω, d) is called a **metric outer measure** if whenever A and B are non-empty subsets of Ω which are positively separated then $\mu(A \cup B) = \mu(A) + \mu(B)$. Their importance lies in the fact that these are precisely the measures on a metric space for which every Borel set is measurable. Metric measures enjoy some very strong continuity properties. Here's one of them.

PROPOSITION 3.0.4. *Let μ be a metric measure on a metric space Ω. Suppose (A_n) is an increasing sequence of subsets of Ω so that A_n and A_{n+1}^c are positively separated. Then*

$$\mu(\cup_n A_n) = \sup_n \mu(A_n).$$

PROOF. We have

$$A_1 \subseteq A_2 \subseteq \cdots \subseteq A_n \subseteq \cdots$$

with A_n and A_{n+1}^c positively separated for each n. We want to show that $\mu(\cup_n A_n) \leq \sup_n \mu(A_n)$ and in this effort we may plainly suppose $\sup_n \mu(A_n) < \infty$ since otherwise all is okay.

First we look at the difference sequence

$$D_1 = A_1, D_2 = A_2 \backslash A_1, D_3 = A_3 \backslash A_2, \ldots, D_n = A_n \backslash A_{n-1}, \ldots$$

Of course, $D_n \subseteq A_n$ but further

$$D_2 \subseteq A_1^c, D_3 \subseteq A_2^c \subseteq A_1^c, D_4 \subseteq A_3^c \subseteq A_2^c \subseteq A_1^c, \ldots$$

So

$$D_1 \subseteq A_1 \text{ and } D_3, D_4, \ldots \subseteq A_2^c,$$
$$D_2 \subseteq A_2 \text{ and } D_4, D_5, \ldots \subseteq A_3^c,$$

etc., etc., etc. In particular

$$\begin{array}{rl}
D_1 \subseteq A_1, & D_3 \subseteq A_2^c \\
D_2 \subseteq A_2, & D_4 \subseteq A_3^c \\
D_1 \cup D_3 \subseteq A_3, & D_5 \subseteq A_4^c \\
D_2 \cup D_4 \subseteq A_4, & D_6 \subseteq A_5^c \\
\vdots & \\
D_1 \cup D_3 \cup \cdots \cup D_{2n-1} \subseteq A_{2n-1}, & D_{2n+1} \subseteq A_{2n}^c \\
D_2 \cup D_4 \cup \cdots \cup D_{2n} \subseteq A_{2n}, & D_{2n} \subseteq A_{2n+1}^c.
\end{array}$$

It follows (inductively if you must know) that

$$\mu(D_1) + \mu(D_3) + \cdots + \mu(D_{2n-1}) + \mu(D_{2n+1}) = \mu(\cup_{k=1}^{n+1} D_{2k-1}),$$

and

$$\mu(D_2) + \mu(D_4) + \cdots + \mu(D_{2n}) + \mu(D_{2n+2}) = \mu(\cup_{k=1}^{n+1} D_{2k}).$$

Each of $\cup_{k=1}^{n+1} D_{2k-1}$ and $\cup_{k=1}^{n+1} D_{2k}$ are subsets of A_{2n+2} and so all above find themselves

$$\leq \mu(A_{2n+2}) \leq \sup_n \mu(A_n) < \infty.$$

Conclusion: both series $\sum_n \mu(D_{2n-1})$ and $\sum_n \mu(D_{2n})$ converge. Now

$$\begin{aligned} \mu(\cup_n A_n) &= \mu(A_n \cup D_{n+1} \cup D_{n+2} \cup \cdots) \quad \text{Not sure about this one.} \\ &\leq \mu(A_n) + \mu(D_{n+1}) + \mu(D_{n+2}) + \cdots \\ &\leq \sup_n \mu(A_n) + \sum_{k=n+1}^{\infty} \mu(D_k); \end{aligned}$$

if $\epsilon > 0$ be given then there is an n so that the latter sum is less than ϵ so

$$\mu(\cup_n A_n) \leq \sup_n \mu(A_n) + \epsilon.$$

Enough said. □

A dividend paid by considering metric outer measures is found in the following.

THEOREM 3.0.5. *If μ is a metric measure on the metric space (Ω, d) then every closed subset of Ω is $\mu-$measurable.* **Consequently every Boreal subset of Ω is $\mu-$measurable.**

Comment that this tells us that Borel sets are $\mu-$measurable.

PROOF. Suppose F is a closed subset of the metric space Ω, and let $A \subseteq F$ and $B \subseteq F^c$ be non-empty sets. For each n let

$$B_n := \left\{ x \in B : \inf_{y \in F} d(x, y) > \frac{1}{n} \right\}.$$

Notice that (B_n) is an ascending sequence of subsets of B with $B = \cup_n B_n$.

By their very definition of B_n, each B_n is positively separated from A. In fact, each B_n is positively separated from B_{n+1}^c. Indeed if $x \in B_n$ and $x' \in B_{n+1}^c$ it's because

$$\inf_{y \in F} d(x, y) > \frac{1}{n}, \quad \text{and} \quad \inf_{y \in F} d(x', y) \leq \frac{1}{n+1}.$$

In the latter situation, it must be that for any $\epsilon > 0$ there is a $y_\epsilon \in F$ so

$$d(x', y_\epsilon) < \frac{1}{n+1} + \epsilon;$$

of course $d(x, y_\epsilon) > \frac{1}{n}$. It follows from the triangle inequality that

$$\begin{aligned} d(x, x') &\geq d(x, y_\epsilon - d(x', y_\epsilon) \\ &> \frac{1}{n} - \left(\frac{1}{n+1} + \epsilon \right) = \frac{1}{n} - \frac{1}{n+1} - \epsilon. \end{aligned}$$

So if we choose $\epsilon_0 = (\frac{1}{n} - \frac{1}{n+1})/2$ we see that

$$d(x, x^{'}) > \left(\frac{1}{n} - \frac{1}{n+1}\right)/2.$$

This we can do for any $x \in B_n$ and $x^{'} \in B_{n+1}^c$.

Let's compute $\mu(A \cup B)$. By our previous theorem

$$\begin{aligned} \mu(A \cup B) &\geq \sup_n \mu(A \cup B_n) \\ &= \sup_n \mu(A) + \mu(B_n) \\ &= \mu(A) + \sup_n \mu(B_n) \\ &= \mu(A) + \mu(\cup_n B_n) \text{ by Proposition } \mathbf{??} \\ &= \mu(A) + \mu(B), \end{aligned}$$

and it follows that F is $\mu-$measurable. □

Since the collection of $\mu-$measurable sets which contain each closed set in Ω is a $\sigma-$fielf, it must be that each Borel set belongs. That Borel sets are $\mu-$measurable whenever μ is a metric outer measure on Ω is **not** accidental; it's part and parcel of being a metric measure. Indeed if we suppose μ is an outer measure on the metric space (Ω, d) for which every closed subset of Ω is $\mu-$measurable and suppose that A and B are positively separated subsets of Ω. Of course $\overline{A} \cap \overline{B} = \emptyset$ so

$$\mu(A \cup B) = \mu((A \cup B) \cap \overline{A}) + \mu((A \cup B) \cap \overline{A}^c)$$

by the measurability of $\overline{A}$. But

$$(A \cup B) \cap \overline{A} = A \text{ and } (A \cup B) \cap (\overline{A})^c = B$$

by the positive separation of A and B. So

$$\mu(A \cup B) = \mu((A \cup B) \cap \overline{A}) + \mu((A \cup B) \cap (\overline{A})^c = \mu(A) + \mu(B),$$

and μ is a metric measure.

4. Haar Measure

Let Q be a fixed compact metric space.

Definition 4.0.6. *We suppose that for subsets of Q the notion of* **congruence** *is defined to satisfy the following conditions (here $A \cong B$ means A is congruent to B):*

(i) $A \cong A$.
(ii) $A \cong B \Leftrightarrow B \cong A$.
(iii) $A \cong B,\ B \cong C \Rightarrow A \cong C$.
(iv) *If A is an open set then so is any set congruent to A.*
(v) *If A is congruent to B and A can be covered by a sequence (A_n) of open sets then B can be covered by a sequence (B_n) so that $B_n \cong A_n$ for each n.*
(vi) *For any open set A the collection of sets congruent to A cover Q.*
(vii) *If (S_n) is a sequence of open concentric balls with radii tending to zero, and if $G_n \cong S_n$ and $a_n, b_n \in G_n$ with $\lim_n a_n$ and $\lim_n b_n$ existing then these limits coincide.*

EXAMPLE 4.0.7. *If Q is a compact metrizable group with left invariant metric then $A \cong B$ whenever $B = xA$ for some $x \in Q$ is a congruence.*

EXAMPLE 4.0.8. *If Q is a compact metric space and G is a group of isometries of Q onto Q that is transitive then $A \cong B$ if $B = \phi(A)$ for some $\phi \in G$ is a congruence.*

Given two relatively compact open sets A, B, by Definition 4.0.6 (vi), the collection of sets congruent to A covers $\overline{B}$, a compact set. Hence there is a finite collection of sets congruent to A that still cover $\overline{B}$. This motivates Haar's covering function $h(B, A)$:

$$h(B, A) = \text{the least number of sets congruent to } A \text{ needed to cover } \overline{B}.$$

PROPOSITION 4.0.9. *Suppose A, B, and C are non-empty open subsets of Q. Then*

(i) $C \subseteq B \Rightarrow h(C, A) \leq h(B, A)$.
(ii) $h(B \cup C, A) \leq h(B, A) + h(C, A)$.
(iii) $B \cong C \Rightarrow h(B, A) = h(C, A)$.
(iv) $h(B, A) \leq h(B, C)h(C, A)$.
(v) *If $d(A, B)$ = distance from A to B is positive (so $\bar{A} \cap \bar{B} = \emptyset$) and (S_n) is sequence of open concentric balls with radii tending to zero, then there is a number N so that for $n \geq N$*

$$h(A \cup B, S_n) = h(A, S_n) + h(B, S_n).$$

PROOF. (v) requires some serious and careful attention. Suppose (v) fails. Then there is (n_k) so that

$$h(A \cup B, S_{n_k}) < h(A, S_{n_k}) + h(B, S_{n_k})$$

for each k. We can plainly suppose that the n_k's are chosen so large that $S_{n_k} \cap A \neq \emptyset$ and $S_{n_k} \cap B \neq \emptyset$ cannot both occur. It follows that there is a sequence (G_k) such that $G_k \cong S_{n_k}$ and $G_k \cap A \neq \emptyset$ and $G_k \cap B \neq \emptyset$. Why? Well fix n_k momentarily and imagine that any G that is congruent to S_{n_k} could meet at most one of A and B. If we cover $A \cup B$ by $h(A \cup B, S_{n_k})$ many sets congruent to S_{n_k}, then this cover (call it $\mathcal{C}$) would be the disjoint union of the collection $\mathcal{A}$ (respectively $\mathcal{B}$) where $\mathcal{A}$ (respectively $\mathcal{B}$) consists of the members of $\mathcal{C}$ that meet only $\mathcal{A}$ (respectively $\mathcal{B}$). Consequently,

$$h(A \cup B, S_{n_k}) = |\mathcal{C}| = |\mathcal{A}| + |\mathcal{B}| \geq h(A, S_{n_k}) + h(B, S_{n_k})$$

which is not an option. So we get a sequence (G_k) of sets with $G_k \cong S_{n_k}$ and $G_k \cap A \neq \emptyset$, $G_k \cap B \neq \emptyset$. From each of the sets $G_k \cap A$ pick a point a_k and from each $G_k \cap B$ pick a point b_k. The sequences $(a_k), (b_k)$ lie inside relatively compact sets so there is a $\mathbb{J} \in P_\infty(\mathbb{N})$ so that

$$a = \lim_{j \in \mathbb{J}} a_j, \quad b = \lim_{j \in \mathbb{J}} b_j$$

both exist. Of course, $a \in \bar{A}, b \in \bar{B}$. But now we're in precisely the position to which (vii) of definition 4.0.6 is applicable: $a_k, b_k \in G_k, G_k \cong S_{n_k}$. Hence $a = b$. But $\overline{A} \cap \overline{B} = \emptyset$. OOPS! The denial of (v) leads to unnecessary chaos. □

Fix a non-empty open subset G of Q. Let (S_n) be a sequence of open concentric balls with radii tending to zero, each contained in G. For any open set $A \subseteq Q$, define

$$l_n(A) = \frac{h(A, S_n)}{h(G, S_n)}.$$

Then

$$h(A, S_n) \leq h(A, G) \cdot h(G, S_n),$$

February 3, 2009

and

$$h(G, S_n) \leq h(G, A) \cdot h(A, S_n),$$

tell us that

$$\frac{1}{h(G, A)} \leq l_n(A) \leq h(A, G).$$

Therefore $(l_n(A))$ is a bounded sequence of real numbers, each of whose terms exceeds the fixed positive number $1/h(G, A)$. Let LIM be a Banach limit, that is,

$$\text{LIM} \in B_{l_\infty^*}$$

and LIM satisfies

$$\liminf x \leq \text{LIM}(x) \leq \limsup x$$

for each and every $x \in l_\infty$. Let

$$l(A) = \text{LIM}((l_n(A))$$

for $A \subseteq G$.

Proposition 4.0.10. *If A and B are open sets then*

(i) $0 < l(A) < \infty$, *as long as* $A \neq \emptyset$.
(ii) $A \subseteq B \Rightarrow l(A) \leq l(B)$.
(iii) $l(A \cup B) \leq l(A) + l(B)$.
(iv) $A \cong B \Rightarrow l(A) = l(B)$.
(v) *If* $d(A, B) > 0$ *then* $l(A \cup B) = l(A) + l(B)$.

Proof. To see (v), note that by Proposition 4.0.9 (v), there exists an N such that for all $n \geq N$,

$$h(A \cup B, S_n) = h(A, S_n) + h(B, S_n).$$

From this we easily see that for all $n \geq N$,

$$l_n(A \cup B) = l_n(A) + l_n(B),$$

and (v) follows.

Let $X \subseteq Q$. Define $\lambda(X)$ as follows:

$$\lambda(X) = \inf \left\{ \sum_n l(A_n) : X \subseteq \bigcup_n A_n,\ A_n \text{ open} \right\}.$$

Here are the fundamental properties of λ.

Theorem 4.0.11. *Let Q be a compact metric space. Then*

(i) $0 \leq \lambda(X)$.
(ii) *If X is a non-empty open subset of Q then* $0 < \lambda(X) < \infty$.
(iii) $X \subseteq Y \subseteq Q \Rightarrow \lambda(X) \leq \lambda(Y)$.
(iv) $X \subseteq \cup_n X_n \Rightarrow \lambda(X) \leq \sum_n \lambda(X_n)$.
(v) $X \cong Y \Rightarrow \lambda(X) = \lambda(Y)$.
(vi) $d(X, Y) > 0 \Rightarrow \lambda(X \cup Y) = \lambda(X) + \lambda(Y)$.

PROOF. (i)-(iv) tells us that λ is a(n) (outer) measure; (v) assures us that λ respects congruence and (vi) says that λ is a 'metric outer measure.' It is a known consequence of λ's metric outer measure character that every Borel set $B \subseteq Q$ is λ-measurable.

(i) deserves comment. If X is open then

$$\lambda(X) \leq l(X) < \infty.$$

If, in addition, $X \neq \emptyset$ then for any $\epsilon > 0$ we can find a sequence (A_n) of open sets so that $X \subseteq \cup_n A_n$ and

$$\lambda(X) + \epsilon \geq \sum_n l(A_n).$$

Now if S is an open ball centered at a point in X and a subset of X, then only finitely many of the A_n's, say $A_1, A_2, \ldots A_N$ are needed to cover S. Then

$$0 < l(S) \leq l(\overline{S}) \leq l(A_1 \cup \ldots A_N) \leq \sum_n l(A_n) < \lambda(X) + \epsilon,$$

and $0 < \lambda(X)$ follows.

(v), too, deserves proof - after all, it's (v) that ensures that λ is a metric outer measure. By Proposition 4.0.10(iv), it suffices to show that $\lambda(X) + \lambda(Y) \leq \lambda(X \cup Y)$. If $d(X, Y) > 0$ then there are disjoint open sets U, V such that $d(U, V) > 0$ with $X \subseteq U$ and $Y \subseteq V$; this is so thanks to normality of metric spaces, if you please. Let $\epsilon > 0$. Pick a sequence (A_n) of relatively compact open sets such that

$$X \cup Y \subseteq \bigcup_n A_n,$$

and

$$\sum_n l(A_n) \leq \lambda(X \cup Y) + \epsilon.$$

Now each of the sets $A_n \cap U, A_n \cap V$ are open and

$$d(A_n \cap U, A_n \cap V) \geq d(U, V) > 0.$$

Hence by Proposition 4.0.10 (v)

$$l(A_n \cap U) \cup (A_n \cap V)) = l(A_n \cap U) + l(A_n \cap V) \leq l(A_n),$$

where the last inequality follows since $A_n \cap U$ and $A_n \cap V$ are disjoint open sets, whose union is a subset of A_n. Further

$$X \subseteq \bigcup_n (A_n \cap U), \;\; Y \subseteq \bigcup_n (A_n \cap V).$$

So

$$\lambda(X) \leq \sum_n l(A_n \cap U), \;\; \lambda(Y) \leq \sum_n l(A_n \cap V).$$

February 3, 2009

This in turns ensures that

$$\begin{aligned}\lambda(X)+\lambda(Y) &\leq \sum_n l(A_n\cap U)+\sum_n l(A_n\cap Y)\\ &\leq \sum_n (l(A_n\cap U)+l(A_n\cap V))\\ &\leq \sum_n l(A_n).\end{aligned}$$

It follows that $\lambda(X)+\lambda(Y)\leq\lambda(X\cup Y)$.

From this we know that λ is a metric outer measure on Q which assigns to any pair of congruent subsets of Q the same value. Hence the collection of $\lambda-$ measurable sets is a $\sigma-$ field of subsets of Q which contains the Borel $\sigma-$ field, and on this $\sigma-$ field, λ is countably additive and assigns congruent measurable sets the same measure.

5. Notes and Remarks

5.1. Saks' Proof of $C(K)^*, K$ a Compact Metric Space. Soon after the appearance of Saks' monograph, Saks published an alternative proof of the theorem of Banach regarding positive linear functionals on $C(Q)$, Q a compact metric space. This proof relies on the theory of metric outer measures to ensure that the resulting measure is a Borel measure.

We set our notation. Let Q be a compact metric space (with metric d), $C(Q)$ is the Banach space of continuous real-valued functions defined on Q, and (to be consistent with Saks) Φ is a positive linear functional on $C(Q)$. Recall that

$$Phi(x)|\leq\Phi(1)$$

whenever $x\in B_{C(Q)}$, so Φ is a member of $C(Q)^*$ with norm $\Phi(1)$.

For any $q\in Q, r>0$ denote by $U_r(q)$ and $B_r(q)$ the sets

$$U_r(q)=\{y\in Q: d(q,y)<r\},\ \ B_r(q)=\{y\in Q: d(q,y)\leq r\}.$$

Stage I For $E\subseteq Q$, define $\lambda(E)$ by

$$\lambda(E):=\inf\{\Phi(x): x\in C(Q), x(q)\geq\chi_E(x),\ \text{ for all } x\in Q\}.$$

Here's what's so about λ:

(i) if $A\subseteq B$ then $\lambda(A)\leq\lambda(B)$;
(ii) if $A,B\subseteq Q$ then $\lambda(A\cup B)\leq\lambda(A)+\lambda(B)$;
(iii) if $d(A,B)>0$ (which is the same as $\overline{A}\cap\overline{B}=\emptyset$) then $\lambda(A\cup B)=\lambda(A)+\lambda(B)$.

(iii) demands comment and proof, even. Let $\epsilon>0$. Pick $x\in C(Q)$ so

$$x(q)\geq\chi_{A\cup B}(q)$$

for all $q\in Q$ (where $A,B\subseteq Q$ satisfy $d(A,B)>0$) and so that

$$\Phi(x)\leq\lambda(A\cup B)+\epsilon.$$

Next chose $h \in C(Q)$ so that $0 \leq h(q) \leq 1$ for all q, with $h(q) = 1$ for $q \in B, h(q) = 0$ for $q \in A$. Let $x_A = (1-h)x$ and $x_B = hx$. Both $x_A, x_B \in C(Q)$; also

$$x_A(q) = \begin{cases} x(q) & \text{if } q \in A \\ 0 & \text{if } q \in B \end{cases} \geq \begin{cases} \chi_{A \cup B}(x) & \text{if } q \in A \\ 0 & \text{if } q \in B \end{cases} = \begin{cases} 1 & \text{if } q \in A \\ 0 & \text{if } q \in B \end{cases} = \chi_A(q),$$

and

$$x_B(q) = \begin{cases} 0 & \text{if } q \in A \\ x(q) & \text{if } q \in B \end{cases} \geq \chi_B(q)$$

for all $x \in Q$. It follows that

$$\begin{aligned} \lambda(A) + \lambda(B) &\leq \Phi(x_A) + \Phi(x_B) \\ &= \Phi(x_A + x_B) \\ &= \Phi(x) \\ &\leq \lambda(A \cup B) + \epsilon. \end{aligned}$$

Stage II Let $E \subseteq Q$ and define $\mu(E)$ by

$$\mu(E) := \inf\{\sum_n \lambda(G_n) : G_n \text{ is open}, E \subseteq \bigcup_n G_n\}.$$

Then μ is an outer measure on Q, a metric outer measure, with the added property that for any closed subset F of Q, $\mu(F) = \lambda(F)$. We'll establish this last claim. Suppose G_n is a sequence of open sets so that

$$F \subseteq \bigcup_n G_n.$$

Since F is closed, Q's compactness is inherited by F, and so we can find N so that

$$F \subseteq G_1 \cup \cdots \cup G_N.$$

It follows that

$$\lambda(F) \leq \sum_{i=1}^{N} \lambda(G_i) \leq \sum_n \lambda(G_n).$$

It follows that

$$\lambda(F) \leq \mu(F).$$

Now let $\epsilon > 0$ be given. Pick $x \in C(Q)$ so that $x(q) \geq \chi_F(q)$ for all q and $\Phi(x) \leq \lambda(F) + \epsilon$. Look at the open set

$$G = \{q \in Q : x(q) > (1 + \epsilon)\}.$$

So $F \subseteq G$ since if $q \in F$ then $x(q) \geq \chi_F(x) = 1$. So

$$\begin{aligned} \mu(F) &\leq \lambda(G) \ \text{(Since } G \text{ is open and } F \subseteq G) \\ &\leq \Phi(x) \ \text{(by definition of } \lambda(G)) \\ &= \Phi(x) \\ &\leq (\lambda(F) + \epsilon). \end{aligned}$$

Let $\epsilon \searrow 0$. Then $\mu(F) \leq \lambda(F)$ follows.

<u>Stage III</u> Like all *metric* outer measures, μ has among its $\mu-$measurable sets each and every Borel

subset of Q. In particular, each $x \in C(Q)$ is $\mu-$measurable, and of course, bounded. Hence each $x \in C(Q)$ is $\mu-$integrable. Let's check $\Phi(x)$ vis-a-vis $\int x d\mu$. We'll show that

$$\Phi(x) \leq \int x\, d\mu$$

for every $x \in C(Q)$.

Take $x \in C(Q)$. Let $\epsilon > 0$. Realizing that $\mu(Q) = \Phi(1)$ we assume that $x(q) > 0$ for all $q \in Q$ - just add an appropriate constant to x and note that this has the exact same effect on $\Phi(x)$ and $\int x d\mu$, leaving their relationship unchanged. Choose $\eta > 0$ so that if $d(q, q^{'}) \leq \eta$ then $|x(q) - x(q^{'}| \leq \epsilon$. Cover Q by open balls $U_{r_1}(q_1), \ldots U_{r_n}(q_n)$, of radii $r_1, \ldots r_n$ each $< \eta/2$ centered at $q_1, \ldots q_n$ respectively, with the added feature that

$$\mu(\{y \in Q : d(y, q_i) = r_i\}) = 0.$$

This last feature can be assumed since $\mu(Q) < \infty$, and so for a fixed $q_0 \in Q$ only countably many of the sets $\{y \in Q : d(y, q_i) = r\}$ can have positive $\mu-$measure. Now that the U's are in place, set

$$\begin{aligned}
E_1 &= \overline{U_{r_1}(q_1)} \\
E_2 &= \overline{U_{r_2}(q_2) \backslash E_1} \\
E_3 &= \overline{U_{r_3}(q_3) \backslash (E_1 \cup E_2)} \\
&\vdots \\
E_n &= \overline{U_{r_n}(q_n) \backslash (E_1 \cup \cdots \cup E_{n-1})}.
\end{aligned}$$

Each set $E_1, \ldots E_n$ is closed with diameter $\leq \eta$ and

$$Q = E_1 \cup \cdots \cup E_n.$$

What's more, and this is crucial, the E_i's overlap only in a set of $\mu-$measure 0! Let $m_i = \min\{x(q) : q \in E_i\},\ i = 1, 2 \ldots n$. Consider

$$a(q) = \sum_{i=1}^{n} m_i \chi_{E_i}(q);$$

notice that

$$x(q) \geq a(q)$$

$\mu-$almost everywhere. Hence

$$\begin{aligned}
\int x\, d\mu &\geq \int \sum_{i=1}^{n} m_i \chi_{E_i}(q)\, d\mu \\
&= \sum_{i=1}^{n} m_i \mu(E_i) = \sum_{i-1}^{n} m_i \lambda(E_i).
\end{aligned}$$

For each $k = 1, \cdots n$, pick $x_k \in C(Q)$ so that for all $q \in Q$, $x_k(q) \geq \chi_{E_k}(q)$ and

$$\mu(E_k) = \lambda(E_k) \geq \Phi(x_k) - \frac{\epsilon}{n m_k}.$$

Put

$$x = m_1 x_1 + \cdots + m_k x_k + \epsilon.$$

Since the oscillation of x on E_k is no more than ϵ then

$$u(q) = \sum_{i=1}^{n} m_i x_i(x) + \epsilon \geq x(q).$$

It follows that

$$\begin{aligned}
\int x\,d\mu &\geq \sum_{i=1}^{n} m_i \mu(E_i) \\
&\geq \sum_{i=1}^{n} m_i \left(\Phi(x_i) - \frac{\epsilon}{n m_i} \right) \\
&= \Phi(\sum_{i=1}^{n} m_i x_i) - \epsilon \\
&= \Phi(u - \epsilon) - \epsilon \\
&= \Phi(u) - \Phi(\epsilon) - \epsilon \\
&\geq \Phi(x) - \epsilon(\Phi(1) + 1).
\end{aligned}$$

Let $\epsilon \searrow 0$ and be happy, don't worry; afterall,

$$\int x d\mu \geq \Phi(x)$$

for all $x \in C(Q)$. $\square$

CHAPTER 6

The Arzela-Ascoli Theorem

CHAPTER 7

Von Neumann's Proof of the Existence and Uniqueness of an Invariant Measure on a Compact Metric group

In this chapter, we'll show how to ascribe to each $f \in C(G)$, a mean $M(f)$, which is at one and the same time, linear in f, non-negative when f is, and is a true average with the values at f and any right translate of f, identical.

Let G be a compact metrizable topological group. Denote by $\mathcal{F}(G)$ the collection of non-empty finite subsets of G and by $C(G)$ the Banach space of all continuous real-valued functions defined on G, equipped as usual with the supremum norm.

Throughout this section, if $F_1, F_2 \in \mathcal{F}(G)$ then by $F_1 \cdot F_2$, we mean all words $a \cdot b$, where $a \in F_1$ and $b \in F_2$; in particular, if $a_1 \cdot b_1 = a_2 \cdot b_2$ but $a_1 \neq a_2$ then we distinguish $a_1 \cdot b_1$ and $a_2 \cdot b_2$.

LEMMA 0.1.1. (i) *If $f \in C(G)$ then $\min f, \max f$, and $Oscf = \max f - \min f$ all exist.*
(ii) *If $f \in C(G)$ and $F \in \mathcal{F}(G)$ then*

$$OscRAve_F f \leq Oscf.$$

In fact,

$$\min f \leq \min RAve_F f \leq \max RAve_F f \leq \max f.$$

(iii) *If $f \in C(G)$ and $F_1, F_2 \in \mathcal{F}(G)$ then*

$$RAve_{F_1} RAve_{F_2} f = RAve_{F_1 \cdot F_2} f.$$

PROOF. To see (ii), let $F \in \mathcal{F}(G)$ and $f \in C(G)$. Define

$$\mathrm{RAve}_{\mathbf{F}} \mathbf{f}(\mathbf{x}) := \frac{\mathbf{1}}{|\mathbf{F}|} \sum_{\mathbf{a} \in \mathbf{F}} \mathbf{f}(\mathbf{xa}), \quad \mathbf{x} \in \mathbf{G}.$$

Naturally $\mathrm{RAve}_F f \in C(G)$.

To see (iii), if $x \in G$ then

$$\begin{aligned}
\mathrm{RAve}_{F_1}\mathrm{RAve}_{F_2} f(x) &= \mathrm{RAve}_{F_1} \frac{1}{|F_2|}\sum_{b\in F_2} f(xb) \\
&= \frac{1}{|F_1|}\sum_{a\in F_1}\frac{S(xa)}{|F_2|} \quad \left(\text{letting } S(x) = \sum_{b\in F_2} f(xb)\right) \\
&= \frac{1}{|F_1|\cdot|F_2|}\sum_{a\in F_1} S(xa) \\
&= \frac{1}{|F_1\cdot F_2|}\sum_{a\in F_1}\sum_{a\in F_2} f(xab) \\
&= \frac{1}{|F_1\cdot F_2|}\sum_{c\in F_1\cdot F_2} f(xc) \\
&= \mathrm{RAve}_{F_1\cdot F_2} f(x). \quad \square
\end{aligned}$$

LEMMA 0.1.2. *If $f \in C(G)$ is not constant then there is an $F \in \mathcal{F}(G)$ such that*

$$OscRAve_F f < Oscf.$$

PROOF. After all, f's not being constant ensures that there is an α such that $\min f < \alpha < \max f$. Set

$$U = [f < \alpha] = \{x \in G : f(x) < \alpha\}.$$

Since $\min f < \alpha$, U is a non-empty open set in G and $G = \bigcup_{a\in G} Ua^{-1}$. (If $x \in G$ then for any $y \in U$, $x = y(y^{-1}x) \in U(y^{-1}x) \subseteq \bigcup_{a\in G} Ua^{-1}$.)

Now U is open (since $f \in C(G)$), and $U \neq \emptyset$ so Ua^{-1} is also a non-empty open set for each $a \in G$. Therefore the Ua^{-1}'s cover the compact G. There is $F \in \mathcal{F}(G)$ such that

$$G = \bigcup_{a\in F} Ua^{-1}.$$

Therefore for any $x \in G$ there exists $a_x \in F$ such that $x \in Ua_x^{-1}$. i.e., for any $x \in G$ there exists $a_x \in F$ such that $f(xa_x) < \alpha$. Thus

$$\begin{aligned}
\mathrm{RAve}_F f(x) &= \frac{1}{|F|}\sum_{a\in F} f(xa) \\
&= \frac{1}{|F|}\left(\sum_{a\in F, a\neq a_x} f(xa) + f(xa_x)\right) \\
&< \frac{1}{|F|}\sum_{a\in F, a\neq a_x} f(xa) + \alpha \\
&\le \frac{(|F|-1)\max f + \alpha}{|F|} \\
&< \frac{(|F|-1)\max f + \max f}{|F|} \\
&= \max f.
\end{aligned}$$

Therefore

$$\text{OscRAve}_F f \leq \text{Osc} f.$$

□

LEMMA 0.1.3. *Let $f \in C(G)$ and define $\mathcal{K} = \{RAve_F f : F \in \mathcal{F}(G)\}$. Then $\mathcal{K}$ is uniformly bounded, equicontinuous family in $C(G)$.*

PROOF. The key to this precious fact is that f is of course uniformly continuous. So given an $\epsilon > 0$ there is an open set V in G containing G's identity such that if $xy^{-1} \in V$ then $|f(x) - f(y)| \leq \epsilon$. Notice that if $a \in G$ and $xy^{-1} \in V$ then $(xa)(ya)^{-1} = xaa^{-1}y^{-1} = xy^{-1} \in V$. So once $xy^{-1} \in V$,

$$|f(xa) - f(ya)| \leq \epsilon$$

for all $a \in G$. But now if $F \in \mathcal{F}(G)$ then whenever $xy^{-1} \in V$ we have

$$\begin{aligned} |\text{RAve}_F f(x) - \text{RAve}_F f(y)| &= \frac{1}{|F|}\left|\sum_{a\in F} f(xa) - \sum_{a\in F} f(ya)\right| \\ &\leq \frac{1}{|F|}\sum_{a\in F}|f(xa) - f(ya)| \\ &\leq \frac{1}{|F|}|F|\epsilon = \epsilon. \end{aligned}$$

Note that $\mathcal{K}$ is uniformly bounded since

$$\begin{aligned} |\text{RAve}_F f(x)| &= \frac{1}{|F|}\left|\sum_{a\in F} f(xa)\right| \\ &\leq \frac{1}{|F|}\sum_{a\in F}|f(xa)| \\ &\leq \frac{1}{|F|}|F| \cdot ||f|| = ||f||_\infty. \quad \square \end{aligned}$$

We see that Lemma 0.1.3 takes on added significance if we but recall the classical theory of Arzela and Ascoli to the effect that $\mathcal{K} \subseteq C(G)$ is relatively norm compact if and only if $\mathcal{K}$ is uniformly bounded and equicontinuous.

With Lemmas 0.1.2 and 0.1.3 in hand, the plan of attack is clear. We want an averaging technique which will give a true average, assigning values in a uniformly distributed manner, If the function f is constant then we will plainly want to assign that value of constancy to f. With the aforementioned lemmas in hand, we handle non-constant functions thusly; if f is not constant, then we can find $F_1 \in \mathcal{F}(G)$ so that

$$\text{OscRAve}_{F_1} < \text{Osc} f;$$

If $\text{RAve}_{F_1}(f)$ is constant then it's value of constancy is the natural value to ascribe to f. If $\text{RAve}_{F_1}(f)$ is not constant, then we appeal to Lemma 0.1.3 again to find $F_2 \in \mathcal{F}(G)$ so that

$$\text{RAve}_{F_2}\text{RAve}_{F_1} \subset \text{OscRAve}_{F_1}(f).$$

Continuing in this vain, we see that in the worst case we can find a sequence $(F_n) \subseteq \mathcal{F}(G)$ so that for each n

$$\text{OscRAve}_{F_{n+1}}\text{RAve}_{F_n}(f) \subset \text{OscRAve}_{F_n}(f).$$

Appealing to Messrs. Arzela and Ascoli, we can pass to a sequence $(F_n^{'}) \subset \mathcal{F}(G)$ so $(\text{RAve}_{F_n'}(f))$ is uniformly convergent.

The point is that because our averages were taken with respect to right translates, in the long run, judicious choices of the F_n's ought to produce an average that is right invariant. Remarkably enough the wisdom needed has already been provided by Von Neumann.

LEMMA 0.1.4. *Let $f \in C(G)$ and $\mathcal{K} = \{RAve_F f : F \in \mathcal{F}(G)\}$. Then*

$$\inf_{g \in \mathcal{K}} Oscg = 0.$$

PROOF. Let

$$s = \inf_{g \in \mathcal{K}} \text{Osc}g = \inf\{\text{OscRAve}_F f : F \in \mathcal{F}(G)\}.$$

Therefore there exists (F_n) in $\mathcal{F}(G)$ such that $(\text{RAve}_{F_n}) \searrow s$. Thanks to Arzela and Ascoli we can, by passing to subsequences if necessary, assume that

$$\text{RAve}_{F_n} f \to g \in C(G),$$

uniformly. It's plain that on assuming the uniform convergence of (RAve_{F_n}) that

$$\min \text{RAve}_{F_n} f \to \min g \text{ and } \max \text{RAve}_{F_n} f \to \max g,$$

and so

$$\text{OscRAve}_{F_n} f \to \text{Osc}g.$$

Thus $\text{Osc}g = s$. Here's the point: g is constant! Indeed if g were not constant there would be an $F_0 \in \mathcal{F}(G)$ such that

$$s_0 = \text{OscRAve}_{F_0} g < \text{Osc}g = s,$$

thanks to lemma 0.1.2. Since $(\text{RAve}_{F_n} f)$ is uniformly convergent, there exists N such that

$$||\text{RAve}_{F_N} f - g||_\infty < \frac{s - s_0}{3}.$$

i.e., for any $x \in G$,

$$|\text{RAve}_{F_N} f(x) - g(x)| \leq \frac{s - s_0}{3}.$$

But this is quickly seen to mean

$$|\text{RAve}_{F_0}\text{RAve}_{F_N} f(x) - \text{RAve}_{F_0} g(x)| \leq \frac{s - s_0}{3},$$

for all $x \in G$ as well. It follows that for all $x \in G$

$$|\text{OscRAve}_{F_0}\text{RAve}_{F_N} f - \text{OscRAve}_{F_0} g| < 2\left(\frac{s - s_0}{3}\right).$$

i.e., for all $x \in G$,

$$|\text{OscRAve}_{F_0}\text{RAve}_{F_N} f(x) - s_0| < 2\left(\frac{s - s_0}{3}\right).$$

But this in turn means that

$$\text{OscRAve}_{F_0}\text{RAve}_{F_N} f(x) < s_0 + 2\left(\frac{s - s_0}{3}\right) = \frac{2}{3}s + \frac{1}{3}s_0 < s.$$

But

$$\text{OscRAve}_{F_0}\text{RAve}_{F_N} f = \text{OscRAve}_{F_0 F_N} f,$$

February 3, 2009

and

$$s = \inf_{F \in \mathcal{F}(G)} \mathrm{Osc}\mathrm{RAve}_F f.$$

This should elicit an 'OOPS' because

$$\mathrm{RAve}_{F_0}\mathrm{RAve}_{F_N} f = \mathrm{RAve}_{F_0 \cdot F_N} f \in \mathcal{K}.$$

Therefore g is constant and $s = 0$. i.e.,

$$\inf_{g \in \mathcal{K}} \mathrm{Osc} g = 0.$$

□

We say the real number p is a **right mean** of f if for each $\epsilon > 0$ there is an $F \in \mathcal{F}(G)$ such that

$$|\mathrm{RAve}_F f(x) - p| < \epsilon$$

for all $x \in G$. i.e.,

$$||\mathrm{RAve}_F f - p||_\infty < \epsilon.$$

THEOREM 0.1.5. *Every $f \in C(G)$ has a right mean.*

PROOF. By the techniques used in Lemma 0.1.4, there is a constant function h (say $h(x) \equiv p$) and a sequence $(F_n) \subseteq \mathcal{F}(G)$ such that

$$\lim_n ||\mathrm{RAve}_{F_n} f - h||_\infty = 0.$$

i.e.,

$$||\mathrm{RAve}_{F_n} f - p||_\infty \to 0,$$

as $n \to \infty$. Plainly p is a right mean of f. □

It's plain that each $f \in C(G)$ has a **left mean** as well, that is, there is a $q \in \mathbb{R}$ so that if $\epsilon > 0$ is given there exists an $F \in \mathcal{F}(G)$ so that

$$\left| \frac{1}{|F|} \sum_{a \in F} f(ax) - q \right| < \epsilon$$

for all $x \in G$. For obvious reasons, we define

$$\mathrm{LAve}_F f(x) = \frac{1}{|F|} \sum_{a \in F} f(ax).$$

THEOREM 0.1.6. *Let $f \in C(G)$. Let p be a right mean of f and q be a left mean of f. Then $p = q$.*

PROOF. Let $\epsilon > 0$. Find $A, B \in \mathcal{F}(G)$ so that

$$||\mathrm{RAve}_A f - p||_\infty \le \frac{\epsilon}{2}, \;\; ||\mathrm{LAve}_B f - q||_\infty \le \frac{\epsilon}{2}.$$

February 3, 2009

Now

$$\begin{aligned}
\mathrm{RAve}_A\mathrm{RAve}_B f(x) &= \mathrm{RAve}_A \frac{1}{|B|}\sum_{b\in B} f(bx) \\
&= \frac{1}{|A|}\frac{1}{|B|}\sum_{a\in A} S(xa) \ \ (\text{where } \ S(x)=\sum_{b\in B} f(bx)) \\
&= \frac{1}{|A|}\frac{1}{|B|}\sum_{a\in A}\sum_{b\in B} f(bxa) \\
&= \frac{1}{|B|}\frac{1}{|A|}\sum_{b\in B}\sum_{a\in A} f(bxa) \\
&= \frac{1}{|B|}\sum_{b\in B}\frac{1}{|A|}\sum_{a\in A} f(bxa) \\
&= \frac{1}{|B|}\sum_{b\in B}\mathrm{RAve}_A f(bx) \\
&= \mathrm{LAve}_B\mathrm{RAve}_A f.
\end{aligned}$$

Further,

$$\mathrm{RAve}_A(\mathrm{LAve}_B f - q) = \mathrm{RAve}_A\mathrm{LAve}_B f - q$$

and

$$\mathrm{LAve}_B(\mathrm{RAve}_A f - p) = \mathrm{LAve}_B\mathrm{RAve}_A f - p.$$

So for any $x \in G$,

$$\begin{aligned}
|p-q| &= |p - \mathrm{RAve}_A\mathrm{LAve}_B f(x) + \mathrm{RAve}_A\mathrm{LAve}_B f(x) - q| \\
&\le |p - \mathrm{RAve}_A\mathrm{LAve}_B f(x)| + |\mathrm{RAve}_A\mathrm{LAve}_B f(x) - q| \\
&= |p - \mathrm{LAve}_B\mathrm{RAve}_A f(x)| + |\mathrm{RAve}_A\mathrm{LAve}_B f(x) - q| \\
&= |\mathrm{LAve}_B(p - \mathrm{RAve}_A f(x))| + |\mathrm{RAve}_A(\mathrm{LAve}_B f(x) - q)| \\
&\le |p - \mathrm{RAve}_A f(x)| + |\mathrm{LAve}_B f(x) - q| \ \ (\text{since } |\mathrm{LAve}_B f| \le |f| \ \text{ and } \mathrm{RAve}_B f| \le |f|) \\
&< \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon,
\end{aligned}$$

and $p = q$. Go figure.

COROLLARY 0.1.7. *For any $f \in C(G)$ there is a unique number $M(f)$ that is both a right and left mean.*

THEOREM 0.1.8. *The functional M on $C(G)$satisfies the following*

(i) *M is linear.*
(ii) *$Mf \geq 0$ if $f \geq 0$.*
(iii) *$M(1) = 1$.*
(iv) *$M({}_af) = M(f) = M(f_a)$ for each $a \in G$, where ${}_af(x) = f(ax)$ and $f_a(x) = f(xa)$.*
(v) *$M(f) > 0$ if $f \geq 0$ but $f \neq 0$.*
(vi) *$M(\hat{f}) = M(f)$ where $\hat{f}(x) = f(x^{-1})$ for each $x \in G$.*

PROOF. We start by showing

$$M(\mathrm{RAve}_F f) = M(f) \tag{0.1}$$

for each $f \in C(G)$ and each $F \in \mathcal{F}(G)$. Suppose that $M(f) = p$. If $\epsilon > 0$ is given to us then we can find $F_0 \in \mathcal{F}(G)$ such that

$$||\text{LAve}F_0 f - p||_\infty \leq \epsilon.$$

i.e.,

$$\left| \frac{1}{|F_0|} \sum_{b \in F_0} f(bx) - p \right| \leq \epsilon$$

for all $x \in G$. It follows that for any $x \in G$ and $a \in F$,

$$|\text{RAve}_F \text{LAve}_{F_0} f(x) - p| \leq \epsilon.$$

Since

$$\text{RAve}_F \text{LAve}_{F_0} f = \text{LAve}_{F_0} \text{RAve}_F f,$$

p is a left mean of $\text{RAve}_F f$. Hence by our previous result,

$$M(\text{RAve}_F f) = p,$$

and

$$M(\text{RAve}_F f) = M(f).$$

To see that M is linear, let $M(f) = p$ and $M(h) = q$. Pick $H_0 \in \mathcal{F}(F)$ so that

$$||\text{RAve}_{H_0} h - q||_\infty \leq \epsilon.$$

i.e., for all $x \in G$,

$$\left| \frac{1}{|H_0|} \sum_{b \in H_0} h(xb) - q \right| \leq \epsilon.$$

i.e., if $E \in \mathcal{F}(G)$ and $x \in G$ then

$$|\text{RAve}_{E \cdot H_0} h(x) - q| = |\text{RAve}_E \text{RAve}_{H_0} h(x) - q| < \epsilon.$$

Now

$$p = M(f) = M(\text{RAve}_F f)$$

for any $F \in \mathcal{F}$. Therefore p is the right mean of $\text{RAve}_{H_0} f$. Hence there exists $F_0 \in \mathcal{F}(G)$ so that

$$||\text{RAve}_{F_0} \text{RAve}_{H_0} f - p|| \leq \epsilon.$$

i.e., for all $x \in G$,

$$|\text{RAve}_{F_0 \cdot H_0} f(x) - p| = |\text{RAve}_{F_0} \text{RAve}_{H_0} f(x) - p| \leq \epsilon.$$

Since we already know that for all $x \in G$ and each $E \in \mathcal{F}(G)$,

$$|\text{RAve}_{E \cdot H_0} h(x) - q| \leq \epsilon,$$

it follows that by taking $E = F_0$ we get for each $x \in G$,

$$|\text{RAve}_{F_0 \cdot H_0} (f + h)(x) - (p + q)| \leq 2\epsilon.$$

Thus

$$M(f + h) = M(f) + M(h).$$

It follows from this and the easily established fact that $M(kf) = kM(f)$ that M is linear, and we have shown (i).

Parts (ii) and (iii) are clear. To see (iv), since

$$\text{RAve}_F f(xa) = \text{RAve}_{a \cdot F} f(x), \tag{0.2}$$

$$\begin{aligned} M(f_a) &= M(\mathrm{RAve}_F f_a(x)) \text{ (by (0.1))} \\ &= M(\mathrm{RAve}_F f(xa)) \\ &= M(\mathrm{RAve}_{a\cdot F} f(x)) \text{ (by (0.2))} \\ &= M(f) \text{ (by (0.1))}. \end{aligned}$$

Similarly,

$$\mathrm{LAve}_F f(ax) = \mathrm{LAve}_{F\cdot a} f(x), \tag{0.3}$$

and so

$$\begin{aligned} M({}_a f) &= M(\mathrm{LAve}_F({}_a f(x))) \text{ (by (0.1) actually it's equivalent with left averages)} \\ &= M(\mathrm{LAve}_F f(ax)) \\ &= M(\mathrm{LAve}_{F\cdot a} f(x)) \text{ (by (0.3))} \\ &= M(f) \text{ (by (0.1) actually it's equivalent with left averages)}, \end{aligned}$$

and thus

$$M({}_a f) = M(f) = M(f_a).$$

For (v), suppose that $f \in C(G), f \geq 0, f \not\equiv 0$. Then there is $\alpha > 0$ such that $U = [f > \alpha]$ is non-empty and open; it's easy to see that $\{U_{a^{-1}} : a \in G\}$ is an open cover of the compact G. (If $x \in G$ then for any $y \in G$, $x = y(y^{-1}x) \in U(y^{-1}x) \subseteq \bigcup_{a\in G} Ua^{-1}$.) It follows that for some $a_1, \ldots, a_m \in G$

$$G = Ua_1^{-1} \bigcup Ua_2^{-1} \bigcup \ldots \bigcup Ua_m^{-1}.$$

Let's check to see how this plays out.

If $x \in G$ then $x \in Ua_k^{-1}$ for some $1 \leq k \leq m$. Hence, $xa_k \in U$ and thus $f(xa_k) > \alpha$. It follows that

$$\mathrm{RAve}_{\{a_1,\ldots,a_m\}} f(x) = \frac{1}{m}\sum_{i=1}^{m} f(xa_i) > \frac{\alpha}{m},$$

for all $x \in G$. Therefore

$$0 < \frac{\alpha}{m} \leq M(\mathrm{RAve}_{\{a_1,\ldots a_m\}} f) = M(f).$$

Almost done; we have but to show that $M(f)$ and $M(\hat{f})$ agree. To establish this, define

$$N(f) = M(f \circ \mathrm{inv}),$$

where $\mathrm{inv} : G \to G$ is given by $\mathrm{inv}(x) = x^{-1}$. N is a linear functional on $C(G)$, $N(f) \geq 0$ if $f \geq 0$, and $N(1) = 1$. Moreover

$$\begin{aligned} N({}_a f) &= M(f_a \circ \mathrm{inv}) \\ &= M({}_{a^{-1}}\hat{f}) \text{ (since } f_a \circ \mathrm{inv}(x) = f_a(x^{-1}) = \hat{f}(a^{-1}x) =_{a^{-1}} f(x)) \\ &= M(\hat{f}) \text{ (by (iv))} \\ &= N(f). \end{aligned}$$

But by Corollary 0.1.7, there is only one invariant mean on $C(G)$ so $N(f) = M(f)$. □

CHAPTER 8

The Fubini-Tonelli Theorem

1. Kakutani's Proof of the Uniqueness of Haar Measure

Let G be a compact topological group. We view G as a group of homeomorphisms of G onto itself. A Borel measure μ is **left $G-$invariant** if for any Borel set $E \subseteq G$ any $x \in G$,

$$\mu(xE) = \mu(E).$$

A Borel measure μ is **right $G-$invariant** if for any Borel set $E \subseteq G$ any $x \in G$,

$$\mu(Ex) = \mu(E).$$

A Borel set $E \subseteq G$ is **left $G - \mu-$invariant** where μ is any Borel measure on G if for each $x \in G$

$$\mu(E \triangle xE) = 0.$$

A Borel set $E \subseteq G$ is **right $G - \mu-$invariant** where μ is any Borel measure on G if for each $x \in G$

$$\mu(E \triangle Ex) = 0.$$

We say that G is **left ergodic** if given a left $G-$invariant countably additive non-negative Borel measure μ then any Borel set $E \subseteq G$ that is left $G - \mu-$invariant satisfies either

$$\mu(E) = 0 \text{ or } \mu(E^c) = 0.$$

Similarly, G is **right ergodic** is defined analogously.

We'll first show that if G is left ergodic then the left invariant measure on G is unique. Then we will show that G is left ergodic.

Assume then that G is left ergodic but μ_1 and μ_2 are both left invariant measures on G for which there are Borel sets $E_1, E_2 \subseteq G$ and a real α such that

$$\mu_1(E_1) < \alpha\mu_2(E_1), \text{ and } \mu_1(E_2) > \alpha\mu_2(E_2)$$

(so that μ_1, μ_2 are not constant multiples of each other). Look at

$$\mu := \mu(E) = \mu_1(E) - \alpha\mu_2(E),$$

a left invariant countably additive Borel measure on G.

The Hahn Decomposition Theorem splits G into the *disjoint* union

$$G = P \cup N$$

of Borel sets P, N in such a way that if E is a Borel subset of P then $\mu(E) \geq 0$ and if E is a Borel subset of N then $\mu(E) \leq 0$. Moreover, P and N are $\mu-$essentially unique in this regard.

February 3, 2009

Now μ is left invariant so P is left invariant and so is N. Why is this so? Let E be a Borel subset of P, and let $y \in G$. Then for any $x \in G$,

$$\mu(xE) = \mu(E) \geq 0.$$

But if F is a Borel subset of yP then $y^{-1}F$ is a Borel subset of $P = y^{-1}yP$ and so

$$\mu(F) = \mu(y^{-1}F) \geq 0.$$

It follows that for any Borel subset F of yP, $\mu(F) \geq 0$ and yP is also a positive set for μ. Thus

$$\mu(P) \geq \mu(yP) \geq 0.$$

Since P is essentially unique as a positive set for μ, it follows that

$$\mu(P) = \mu(yP).$$

In a similar fashion we see that N is left $\mu-$invariant.

Look at the countably additive, non-negative, left invariant measure $|\mu|$, the variation of μ,

$$|\mu|(E) = \mu(E \cap P) + \mu(E \cap N).$$

Ergodicity of G says

$$|\mu|(P) = 0 \text{ or } |\mu|(N) = 0,$$

since each of P and N is $G - |\mu|-$invariant. *But* $\mu(E_1) > 0$ and $\mu(E_2) < 0$. Therefore

$$|\mu|(P) \geq \mu(E_1 \cap P) \geq \mu(E_1) > 0, \text{ and } |\mu|(N) \geq -\mu(E_2 \cap N) \geq -\mu(E_2) > 0,$$

Oops! The uniqueness of left-invariant measures on G follows.

We now show that G is left ergodic. Let μ be a left invariant countably additive non-negative (real-valued) Borel measure on G and let E be a left $\mu-$invariant Borel subset of G.

Look at $\mu \otimes \lambda$ on $G \times G$ where λ is an arbitrary but fixed *right* invariant regular Borel measure on G. Consider χ_E, a Borel function on G to be sure. Set

$$\phi(x, y) = \chi_E(yx).$$

Since χ_E is a Borel measurable function ϕ is a Borel measurable function on $G \times G$. For any $y \in G$,

$$|\chi_E(x) - \phi(x, y)| = \chi_{E \triangle y^{-1}E}(x).$$

But our assumption on E is that E is left $\mu-$invariant; hence

$$\mu(E \triangle yE) = 0$$

for each $y \in G$; in particular,

$$\int |\chi_E(x) - \phi(x, y)| \, d\mu(x) = \int \chi_{E \triangle y^{-1}E}(x) \, d\mu(x) = \mu(E \triangle y^{-1}E) = 0,$$

for each $y \in G$. It follows that

$$\int_G \int_G |\chi_E(x) - \phi(x, y)| \, d\mu(x) \, d\lambda(y) = 0.$$

Monsieur Fubini steps in to say that for $\mu-$almost all $x \in G$, we have

$$\int |\chi_E(x) - \phi(x, y)| \, d\lambda(y) = 0.$$

He further stipulates that there exists $M \subseteq (G, \mu)$ M a Borel set, $\mu(M) = 0$ so if $x \notin M$ there exists $C_x \subseteq (G, \lambda)$, C_x a Borel set, $\lambda(C_x) = 0$ so for any $y \notin C_x$ we have $\chi_E(x) = \chi_E(yx)$. To see this, take $x \notin M$. Then $y \notin C_x$ means $yx \notin C_x \cdot x$; on letting $z = yx$ we have for $z \notin C_x \cdot x$ that $\chi_E(x) = \chi_E(z)$.

Now look at $x^{'} \in M$. Then there exists $C_{x'} \subseteq (G, \lambda)$, $C_{x'}$ a Borel set, $\lambda(C_{x'}) = 0$ so if $y \notin C_{x'}$ then $\chi_E(x^{'}) = \chi_E(yx^{'})$. Again, $y \notin C_{x'}$ is the same as $yx^{'} \in C_{x'} \cdot x^{'}$ so if $z = yx^{'}$ we have $z \notin C_{x'} \cdot x^{'}$, implying $\chi_E(x^{'}) = \chi_E(z)$.

Since

$$\lambda(C_x \cdot x) = \lambda(C_x) = 0 = \lambda(C_{x'} \cdot x^{'}) = \lambda(C_{x'} \cdot x^{'}),$$

$C_x \cdot x, C_{x'} \cdot x^{'}$ both have the same $\lambda-$measure as C_x and $C_{x'}$, respectively; that is,

$$\lambda(C_x \cdot x) = 0 = \lambda(C_{x'} \cdot x^{'}).$$

So there exists $z \in G \backslash ((C_x \cdot x) \cup (C_{x'} \cdot x^{'}))$. For such a z,

$$\chi_E(x) = \chi_E(z) = \chi_E(x^{'}).$$

In other words χ_E is constant on M^c. Check out the possibilities:

$$\chi_E = 0 \text{ or } \chi_E = 1.$$

Either way is okay!

CHAPTER 9

Homogeneous Spaces

Let G be a compact topological group and K be a compact Hausdorff space. We say that G **acts transitively** on K if there is a continuous map $G \times K \to K : (g, k) \mapsto g(k)$ such that

(i) $e(k) = k$ for all $k \in K$ (e is the identity of G);
(ii) $(g_1 g_2)(k) = g_1(g_2(k))$ for all $g_1, g_2 \in G, h \in K$;
(iii) given $k_1, k_2 \in K$ there is a $g \in G$ so that $g(k_1) = k_2$.

It is noteworthy that each $g \in G$ may be viewed as a homeomorphism of K onto itself; after all, the map $k \to g(k)$ is continuous and has $k \to g^{-1}(k)$ as an inverse.

Condition (iii) says, in particular, that the space K is *homogeneous*;i.e., we can move points of K around K via homeomorphisms (members of G, in fact) of K onto itself.

If μ is the unique translation invariant Borel probability on G then μ induces a $G-$invariant Borel probability on K. This is an important construction, one worth understanding in general as well as in special cases. **WE HAVE NOT USED THE WORD 'PROBABILITY' ANYWHERE BEFORE THIS PARAGRAPH.**

Suppose H is a closed subgroup of the compact topological group G. Consider the set G/H with the so-called 'quotient topology,' that is, the strongest topology that makes the natural map $q_H : G \to G/H$ (taking $g \in G$ to $gH \in G/H$) continuous; so $U \subseteq G/H$ is open precisely when $q_H^{\leftarrow}(U)$ is open in G. In other words, a typical open set in G/H is of the form $\{xH : x \in V\}$ when V is open in G. Because H is supposed to be closed, this topology is Hausdorff; because q_H is continuous and surjective, G/H is compact.

More is so. G acts transitively on G/H. The map $(g, g^{'}H) \to gg^{'}H$ fits the bill in the definition.

In fact, any transitive action of a compact group on a compact space is of the sort just described. To be sure we need to tell when seemingly different spaces are the same under G's action. Let G act transitively on each of the compact Hausdorff spaces K_1, K_2. We say that K_1 and K_2 are **isomorphic under G's action** if there is a homeomorphism $\phi : K_1 \to K_2$ such that

$$\phi(g(k_1)) = g(\phi(k_1))$$

for each $k_1 \in K_1$.

Theorem 0.0.9 (Weil). *Let the compact group G act transitively on the compact Hausdorff space K. Then there is a closed subgroup H of G such that K and G/H are isomorphic under G's action.*

PROOF. Fix $k_0 \in K$. Look at

$$H = \{g \in G : g(k_0) = k_0\}.$$

H is called the **isotopy subgroup**. It is plain that H is a closed subgroup of G. A natural candidate for the isomorphism of G/H and K is at hand: $\phi : G/H \to K$ given by

$$\phi(gH) = g(k_0).$$

For $g_1, g_2 \in G$, $g_1(k_0) = g_2(k_0)$ precisely when

$$g_1^{-1}(g_2(k_0)) = g_1^{-1}(g_1(k_0)) = e(k_0) = k_0,$$

or $g_1^{-1}g_2 \in H$, which is tantamount to $g_1 H = g_2 H$. This assures us that ϕ is well-defined and injective.

The transitivity of G's action ensures ϕ's surjectivity. To see that ϕ is also continuous, fix $g \in G$ and let V be an open set in K containing $g(k_0)$. By the continuity of the map

$$(g, k) \to g(k)$$

on $G \times K$, there is an open set U in G which contains g so that $u(k_0) \in V$ for all $u \in U$. But $q_H(U)$ is open in G/H's quotient topology and $q_H(U) \subseteq \phi^{\leftarrow}(V)$. This shows that ϕ is a continuous bijection between the compact Hausdorff spaces G/H and K; as such ϕ is a homeomorphism.

Further if $g_1, g_2 \in G$ then

$$g_1(\phi(g_2 H)) = g_1(g_2(k_0)) = (g_1 g_2)(k_0) = \phi(g_1(g_2 H)).$$

Thus G/H and K are isomorphic under G's action. $\square$

Note that because of this isomorphism theorem we can consider *any* G/H where H is the isotopy subgroup associated with *any* $k_0 \in K$.

Now we're ready for the main course.

THEOREM 0.0.10 (Weil). *Suppose the compact group G acts transitively on the compact Hausdorff space K. Then there is a unique $G-$invariant regular Borel probability measure on K.*

PROOF. We identify K with the isotopy subgroup G/H as in our previous theorem. Let

$$q_H : G \to G/H$$

be the natural quotient map. Suppose μ is the normalized Haar measure on G and define $\mu_{G/H}$ on G/H by

$$\mu_{G/H}(B) = \mu(q_H^{\leftarrow}(B))$$

for any Borel set $B \subseteq G/H$.

If $g \in G$ and B is Borel subset of G/H then

$$\begin{aligned} g(q_H^{\leftarrow}(B)) &= \{gx : xH \in B\} \\ &= \{gx : gxH \in gB\} \\ &= q_H^{\leftarrow}(gB). \end{aligned}$$

Therefore

$$\begin{aligned}\mu_{G/H}(gB) &= \mu(q_H^{\leftarrow}(gB))\\ &= \mu(g(q_H^{\leftarrow}(B))\\ &= \mu(q_H^{\leftarrow}(B)) = \mu_{G/H}(B),\end{aligned}$$

and $\mu_{G/H}$ is $G-$invariant.

Uniqueness is a touchier issue, as is always the case it seems. We take a close look at how members of rca($\mathcal{B}_0(G)$) act on $C(G)$. Take $\phi \in C(G)$ and $g \in G$. **DEFINE rca SOMEWHERE?** Define $\phi_g \in C(G)$ by

$$\phi_g(x) = \phi(gx).$$

Denote by μ_H the Haar measure (normalized so $\mu_H = 1$) on H. The map $G \to C(G)$ that takes g to ϕ_g is uniformly continuous (this is an easy modification of Theorem 3.0.9) so that

$$\hat{\phi}(g) = \int_H \phi_g(h) d\mu_H(h), \ \ g \in G$$

defines a member $\hat{\phi}$ of $C(G)$.

Suppose $g_1H = g_2H$. Then $g_1^{-1}g_2 \in H$,

$$\begin{aligned}\hat{\phi}(g_1) &= \int_H \phi_{g_1}(h)\, d\mu_H(h)\\ &= \int_H \phi_{g_1}(g_1^{-1}g_2h)\, d\mu_H(h) \ \ (\text{by } \ \mu_H\text{'s invariance and } \ g_1^{-1}g_2 \in H)\\ &= \int_H \phi(g_2H)\, d\mu_H(h)\\ &= \int_H \phi_{g_2}(h)\, d\mu_H(h) = \hat{\phi}(g_2).\end{aligned}$$

Therefore $\hat{\phi}$ is constant on the left cosets of H so we can lift $\hat{\phi}$ to a continuous function $\tilde{\phi}$ on G/H:

$$\tilde{\phi}(gH) = \hat{\phi}(g).$$

To summarize: if $\phi \in C(G)$ then we define $\hat{\phi} \in C(G)$ and from this we get $\tilde{\phi} \in C(G/H)$. Remarkably, each member of $C(G/H)$ comes about from this procedure. In fact, if $f \in C(G/H)$ then $f \circ q_H \in C(G)$ and for any $g \in G$

$$\begin{aligned}(\widetilde{f \circ q_H})(gH) &= (\widehat{f \circ q_H})(g)\\ &= \int_H (f \circ q_H)_g(h)\, d\mu_H(h)\\ &= \int_H (f \circ q_H)(gh) d\mu_H(h)\\ &= \int_H f(ghH) d\mu_H(h)\\ &= \int_H f(gH) d\mu_H(h)\\ &= f(gH)\mu_H(H) = f(gH).\end{aligned}$$

In other words, $f = (\widetilde{f \circ q_H})$.

Now we look at G's action. Take any $G-$invariant regular Borel probability measure ν on G/H. For $\phi \in C(G)$ define

$$\lambda(\phi) = \int_{G/H} \tilde{\phi}(gH)\, d\nu(gH).$$

Then λ is a probability measure in $C(G)^\star$. Moreover, λ is translation invariant. Indeed if $x \in G$

$$\begin{aligned}
\lambda(\phi_x) &= \int_{G/H} \tilde{\phi}_x(gH)\, d\nu(gH) \\
&= \int_{G/H} \hat{\phi}_x(g)\, d\nu(gH) \\
&= \int_{G/H} \phi(xg)\, d\nu(gH) \\
&= \int_{G/H} \tilde{\phi}(xgH)\, d\nu(gH) \\
&= \int_{G/H} \tilde{\phi}(gH)\, d\nu(gH) = \lambda(\phi).
\end{aligned}$$

So λ is nothing else but normalized Haar measure on G. **WE SHOULD TALK ABOUT WHAT IT MEANS TO BE 'TRANSLATION INVARIANT' IN TERMS OF MEMBERS OF** $C(K)$.

If ν_1 and ν_2 are $G-$invariant regular Borel probabilities on G/H and if $x = (\widetilde{x \circ q_H}) \in C(G/H)$ then

$$\nu_1(x) = \lambda(x \circ q_H) = \nu_2(x);$$

in other words, ν_1 and ν_2 are the same. **WHY?** □

The worth of an abstract construction lies, at least in part, it its applicability to concrete cases. Our first application is classical and was well-known before Weil's general theorem. It is, nonetheless, interesting and important.

Our setting: $O(n)$, the orthogonal group of order n is our compact group; S^{n-1}, the unit sphere in $\mathbb{R}^n$ is our compact Hausdorff space. The action of $O(n)$ on S^{n-1} is given, naturally by

$$(u, x) \rightarrow u(x).$$

It is easy to verify that $O(n)$ acts on S^{n-1} in a suitable fashion! Transitivity follows by letting $x, x^{'} \in S^{n-1}$; choose orthonormal bases $x_1, x_2, \ldots, x_n$ and $x_1^{'}, x_2^{'}, \ldots, x_n^{'}$ for $\mathbb{R}^n$, and let $u : \mathbb{R}^n \rightarrow \mathbb{R}^n$ be the member of $O(n)$ that takes x to $x^{'}$, x_j to $x_j^{'}$ for $j = 2, \ldots, n$.

Acknowledging the descriptions of members of $O(n)$ as rotations of $\mathbb{R}^n$, a direct application of Weil's theorem says: *There is a unique rotation-invariant regular Borel probability measure on* S^{n-1}.

Geometry is replete with examples of compact Hausdorff spaces that are homogeneous spaces on which various compact groups act transitively.

Here are a few more.

Again our group is $O(n)$. This time our underlying compact space is

$$\sum(n) = \{(x,y) \in S^{n-1} \times S^{n-1} : x \perp y\},$$

where $x \perp y$ means x is perpendicular to y. Note that $(x,y) \in \sum(n)$ precisely when for any real valued numbers a, b:

$$||ax+by||^2 = a^2 + b^2.$$

It is easy to see from this that $\sum(n)$ is a closed subset of $S^{n-1} \times S^{n-1}$, hence, is compact. The action of $O(n)$ is natural enough, too: $(u,(x,y)) \to (ux, uy)$. It is quick and reasonably painless to see that $O(n)$ acts transitively on $\sum(n)$.

One more. Let $1 \le m \le n$. Denote by $\sum^{(m)}(n)$ the set

$$\sum^{(m)}(n) = \left\{ (x_1, \dots, x_m) \in \underbrace{S^{n-1} \times \cdots \times S^{n-1}}_{m \text{ times}} : \{x_1, \dots, x_m\} \text{ is orthonormal} \right\}.$$

Note that $(x_1, \dots, x_m) \in \sum^{(m)}(n)$ precisely when regardless of the real numbers $a_1, \dots a_m$, we have

$$\left|\left| \sum_{j=1}^m a_j x_j \right|\right|^2 = \sum_{j=1}^m a_j^2.$$

This in mind, $\sum^m(n)$ is a compact set of $(S^{n-1})^m$ is easy to see; moreover, the action of $O(n)$ on $\sum^{(m)}(n)$ is given by

$$(u,(x_1, \dots, x_m)) \to (ux_1, ux_2, \dots ux_m)$$

is a transitive one, establishing, with a modicum of tender love and care, that $O(n)$ acts transitively on $\sum^{(m)}(n)$.

Next let $\mathcal{G}_m(n)$ denote the $m-$dimensional Grassmanian manifold, that is, $\mathcal{G}_m(n)$ is the space of all $m-$dimensional linear subspaces of $\mathbb{R}^n$. There is a natural surjection of $\sum_{(m)}(n)$ onto $\mathcal{G}_m(n)$ that takes $(x_1, \dots, x_m) \in \sum^{(m)}(n)$ to the linear span of $\{x_1, \dots, x_m\} \in \mathcal{G}_m(n)$. If we equip $\mathcal{G}_m(n)$ with the natural quotient topology the result is a compact Hausdorff space. Clearly $O(n)$ acts transitively on $\mathcal{G}_m(n)$. The map reflecting the action of $O(n)$ on $\mathcal{G}_m$ is plain: if $\{x_1, \dots, x_m\}$ is an orthonormal set in $\mathbb{R}^n$ then

$$(u, \text{span}\{x_1, \dots, x_m\}) = \text{span}\{ux_1, \dots, ux_m\}.$$

Here we interject that the geometry imparted above on $\mathcal{G}_m(n)$ is such that if $E = \text{span}\{x_1, \dots, x_m\}$ and $E^{'} = \text{span}\{x_1^{'}, \dots, x_m^{'}\}$ are members of $\mathcal{G}_m(n)$ and if each x_k is close to $x_k^{'}$ in $\mathbb{R}^n$ then E is close to $E^{'}$ in $\mathcal{G}_m(n)$.

In this way we find that there is a unique rotation invariant probability Borel measure on the $n-$dimensional Grassmanian manifold $\mathcal{G}_m(n)$.

CHAPTER 10

Metrics in Compact Groups

Let G be a locally compact metrizable group with left invariant metric. Then G has a neighborhood basis of open balls with compact closure. Can G have a metric in which all of the balls have compact closure? If so then G must be second countable; afterall $G = \cup_n B_n$ where the B_n's are centered at a fixed point of G and have a radius n. Since each B_n has $\overline{B}_n$ compact, and since G is separable and metrizable, it's second countable.

Here's a theorem of R. Struble. [**?**]

THEOREM 0.0.11. *A locally compact group metrizable topological group has a left invariant metric (that generates its topology) in which all its open balls have compact closure if and only if G is second countable.*

Though the first Lemma's content is a consequence of the Birkhoff-Kakutani thereom, the proof rendered here (and due to Struble) is too clever and enlightening not to be included.

LEMMA 0.0.12. *Let G be a locally compact group with left Haar measure. Let (V_n) be a decreasing sequence of open sets that form a neighborhood basis of the identity e in G where $\overline{V}_n$ compact for each n. Then*

$$\rho(x,y) = \sup_n \mu(xV_n \triangle yV_n)$$

defines a left invariant metric on G which is compatible with the topology of G.

PROOF. It's clear that $\rho(x,y)$ is well-defined and that $\rho(x,y) = \rho(y,x)$. Moreover $\rho(x,y) \geq 0$ and $\rho(x,y) < \infty$ regardless of $x, y \in G$ since each V_n is a Borel set with compact closure. Further, $\rho(zx,zy)$ and $\rho(x,y)$ coincide because μ is left invariant.

If $x \neq y$ then since G is Hausdorff there must be an m so that $xV_m \cap yV_m = \emptyset$; but now

$$\rho(x,y) \geq \mu(xV_m \triangle yV_m) = 2\mu(V_m) > 0.$$

On noticing that for any n and any $x, y, z \in G$,

$$xV_m \triangle yV_n \subseteq (xV_n \triangle zV_n) \cup (zV_n \triangle yV_n),$$

we see that for any $x, y, z \in G$,

$$\begin{aligned} \mu(xV_n \triangle yV_n) &\leq \mu((xV_n \triangle zV_n) \cup (zV_n \triangle yV_n)) \\ &\leq \mu(xV_n \triangle zV_n) + \mu(zV_n \triangle yV_n) \\ &\leq \rho(x,z) + \rho(z,y), \end{aligned}$$

and with this

$$\rho(x,y) \leq \rho(x,z) + \rho(z,y).$$

In sum, ρ is a left invariant metric on G. If G's topology is discrete then $\mu(\{e\}) > 0$ so $V_m = \{e\}$ for some m; hence if $x \neq y$,

$$\rho(x,y) \geq \mu(xV_m \triangle yV_m) = \mu(\{x,y\}) = 2\mu(\{e\}) > 0,$$

and the topoology induced by ρ is discrete. If G's topology is not discrete then $\mu(V_n) \searrow \mu(\cap_n V_n) = \mu(\{e\}) = 0$. If V is any open set containing e then there is an $m \in \mathbb{N}$ so that $V_m V_m^{-1} \subseteq V$.

Claim 1: $x \in V$ whenever $\rho(x,e) < \mu(V_m)$. To see this, let $\rho(x,e) < \mu(V_m)$. Then

$$\mu(xV_m \triangle V_m) \leq \rho(x,e) < \mu(V_m),$$

a positive number. Were $xV_m \cap V_m = \emptyset$ then

$$\mu(xV_m \triangle V_m) = 2\mu(V_m) < \mu(V_m),$$

oops! So $xV_m \cap V_m \neq \emptyset$ and thus there are $v_1, v_2 \in V_m$ so $xv_1 = v_2 \in xV_m \cap V_m$ and

$$x = v_2 v_2^{-1} \in V_m V_m^{-1} \subseteq V.$$

This is so whenever $\rho(x,e) < \mu(V_m)$, and our claim is justified.

Let's look at all of the points x such that $\rho(x,e) < r$, where $r \in \mathbb{Q}, r > 0$. There must be an $m \in \mathbb{N}$ so that $\mu(V_n) < \frac{r}{4}$, whenever $n \geq m$. Each of the functions

$$f_k(x) = \mu(xV_k \triangle V_k)$$

is continuous and satisfies $f_k(e) = \mu(V_k \triangle V_k) = \mu(\emptyset) = 0$. But now we know there is an $l \in \mathbb{N}$ so that if $x \in V_l$ then $f_1(x), \ldots, f_{m-1}(x) < r$.

Claim 2: if $x \in V_l$ then $\rho(x,e) < r$. Why is this so? Well if $x \in V_l$ then by choice of $l \in \mathbb{N}$, we have

$$\mu(xV_1 \triangle V_1), \ldots, \mu(xV_{m-1} \triangle V_{m-1}) < r.$$

What about $\mu(xV_k \triangle V_k)$ for $k \geq m$? In this case,

$$\mu(xV_k \triangle V_k) \leq 2\mu(V_k) < 2 \cdot \frac{r}{4} < r.$$

It follows that $\rho(x,e) = \sup_n \mu(xV_m \triangle V_n) < r$.

Our two claims taken in tandem show that ρ generates G's topology about e. Since ρ is left invariant and since G's topology is too this is enough to say that ρ generates G's topology everywhere. $\square$

LEMMA 0.0.13. *Let G be a locally compact, second countable (hence metrizable, separable) group. Then there exists a family $\{U_r : r \in \mathbb{N}, r > 0\}$ such that*

(i) for each r, each U_r is open and $\overline{U}_r$ is compact,
(ii) $U_r = U_r^{-1}$
(iii) $U_r U_s \subseteq U_{r+s}$ (so if $r < s$ then $U_r \subseteq U_r U_{r-s} \subseteq U_s$),
(iv) $\{U_r : r > 0\}$ is a base for the open sets about e,
(v) $\cup_{r>0} U_r = G$.

February 3, 2009

Once Lemma 0.0.13 is established, we're ready for business. Indeed, let $\{U_r : r > 0\}$ be the family of open sets about e generated from Lemma 0.0.13. For $x, y \in G$, set

$$d(x, y) = \inf\{r : y^{-1}x \in U_r\}.$$

- Since $G = \cup_{r>0} U_r$, for an pair $x, y \in G$, we have $y^{-1}x \in U_r$ for some $r > 0$. It follows that $d(x, y) \geq 0$.
- $e \in U_r$ for each $r > 0$ so $d(x, x) = 0$.
- If $y^{-1}x \neq e$ then there is an $r_0 > 0$ so that $y^{-1}x \notin U_{r_0}$ (Part (iv) of Lemma 0.0.13) tells us this). But whenever $0 < r_0 < r$ we have (by Part (iii) of Lemma 0.0.13)

$$U_{r_0} \subseteq U_{r_0} U_{r-r_0} \subseteq U_r,$$

  so $d(x, y) \geq r_0 > 0$.
- $U_r = U_r^{-1}$ so $y^{-1}x \in U_r$ precisely when $x^{-1}y \in U_r$; consequently, $d(x, y) = d(y, x)$.
- Suppose $x, y, z \in G$ with $y^{-1}x \in U_r$, $z^{-1}y \in U_s$. Then

$$z^{-1}x = z^{-1}yy^{-1}x \in U_s U_r \subseteq U_{r+s},$$

  so $d(x, z) \leq r + s$. This is so whenever $y^{-1}x \in U_r$ so $d(x, z) \leq d(x, y) + s$; again this is so whenever $z^{-1}y \in U_s$ so $d(x, z) \leq d(x, y) + d(y, z)$.
- Finally, if $x, y, z \in G$ then

$$d(zx, zy) = \inf\{r : (zy)^{-1}zx \in U_r\} = \inf\{r : y^{-1}x \in U_r\} = d(x, y).$$

To summarize: d is a left invariant metric on G.

Since $d(x, e) < r$ means $x = e^{-1}x \in U_r$, the open d-ball of radius r centered at e is contained in U_r. Also this same d-ball contains $U_{r'}$ for any $0 < r' < r$ since if $x \in U_{r'}$ then

$$e^{-1}x = x \in U_{r'} \subseteq U_{r'} U_{\frac{r-r'}{2}} \subseteq U_{\frac{r+r'}{2}},$$

and so $d(x, e) \leq \frac{r+r'}{2} < r$. Therefore if $0 < r' < r$ then

$$U_{r'} \subseteq \{x : d(x, e) < r\} \subseteq U_r.$$

Thus by Part (iv) of Lemma 0.0.13 the metric d is compatible with the topology of G and by Part (i) of Lemma 0.0.13, all d-balls are bounded. Joe: you wrote the following instead but I had troubles reading your handwriting: Thus the open d-balls of radius r about e are ??? with the collection $\{U_r : r > 0\}$, so the closure of each open d-ball is compact

Proof. (Lemma 0.0.13) Let ρ be the left invariant metric resulting from Lemma 0.0.12. We can assume that each of the open balls

$$B_r = \{x \in G : \rho(x, e), r\}$$

has compact closure for $0 < r \leq 2$; afterall, there is an r_0 so that for $r < r_0$, $\overline{B}_{r_0}$ is compact by G's locally compact nature so recalibrate ρ to make $r_0 = 2$ if necessary.

For $0 < r < 2$ we let $U_r = B_r$. This assures us clearly of (iv) and since we'll keep these U_r's, (iv) is assumed henceforth. Also (i), (ii), and (iii) hold when $r + s < 2$ by ρ's left invariant metric nature.

G is locally compact and satisfies the second countability axiom so G admits a countable open base

$$\{W_{2^n} : n \in \mathbb{N}\}$$

for its topology, where we can (and do) assume that $\overline{W}_{2^n}$ is compact for each n. We define

$$U_2 = B_2 \cap W_2.$$

It's easy to verify that (i) and (ii) hold for $0 < r < 2$ and if $r + s < 2$ then (iii) holds as well.

We'll now inch our way from from (i), (ii), and (iii), $(r+s < 2)$ holding for $0 < r \leq 2$ to $0 <\leq r \leq 4$.

First we have to define U_r for $2 < r < 2^2$. Let $0 < r < 2^2$. Set

$$U_r = \bigcup U_{t_1} \cdots U_{t_m}$$

where the union extends over all $t_1, \ldots t_m$ so that each t_i satisfies $0 < t_i \leq 2$ and $t_1 + \cdots + t_m = r$.

If $2 < r < 2^2$ and $t_1 + \cdots t_m = r$ where each $t_i > 0$ then there must be $k, l \in \mathbb{N}$ so that $1 \leq k < l < m$ and $t_1 + \cdots t_k \leq 2$, $t_{k+1} + \cdots + t_l \leq 2$, and $t_{l+1} + \cdots t_m \leq 2$. Why is this so? Well let k be the least j_1 so that $t_1 + \cdots t_{j_1} \leq 2$, and let l be the least j_2 so that $t_{j_1+1} + \cdots t_{j_2} \leq 2$. Then $\sum_{j_2+1}^{m} t_j \leq 2$ because otherwise, $t_{j_k+1} + \cdots t_m > 2$ and $t_1 + \cdots t_{j_1+1} \geq 2$ too where $j_1 + 1 < j_2 + 1$.

It follows that

$$\begin{aligned} U_{t_1} \cdot U_{t_2} \cdots U_{t_m} &\subseteq (U_{t_1} \cdots U_{t_k})(U_{t_{k+1}} \cdots U_{t_l})(U_{t_{l+1}} \cdots U_{t_m}) \\ &\subseteq U_{t_1+\cdots+t_k} U_{t_{k+1}+\cdots+t_l} U_{t_{l+1}+\cdots+t_m} \quad \text{by (iv)} \\ &\subseteq U_2 \cdot U_2 \cdot U_2, \end{aligned}$$

so $U_r \subseteq U_2 \cdot U_2 \cdot U_2$ whenever $0 < r < 2^2$. $\overline{U}_2$ is compact so $\overline{U_2 \cdot U_2 \cdot U_2} \subseteq \overline{U}_2 \cdot \overline{U}_2 \cdot \overline{U}_2$ is too and $\overline{U}_r$ is compact for $0 < r < 2^2$. Since

$$(U_{t_1} \cdots U_{t_m})^{-1} = U_{t_m}^{-1} \cdots U_{t_1}^{-1} = U_{t_m} \cdots U_{t_1},$$

we see that