INVARIANT MEASURES

JOHN VON NEUMANN

AMERICAN MATHEMATICAL SOCIETY
Providence, Rhode Island

2000 *Mathematics Subject Classification.* Primary 22Axx, 22A10.

Library of Congress Cataloging-in-Publication Data

Von Neumann, John, 1903–1957.
Invariant measures / John von Neumann.
p. cm.
Includes index.
ISBN 0-8218-0912-1 (alk. paper)
1. Invariant measures. I. Title.
QA325.V66 1998
515′.42—dc21 98-15971
CIP

♾ The paper used in this book is acid-free and falls within the guidelines
established to ensure permanence and durability.
Visit the AMS home page at URL: `http://www.ams.org/`

10 9 8 7 6 5 4 3 2 04 03 02 01 00

Contents

Preface

In 1940–1941 von Neumann lectured on invariant measures at the Institute for Advanced Study. This book is essentially a written version of what he said.

The lectures began with general measure theory and went on to Haar measure and some of its generalizations. Shizuo Kakutani was at the Institute that year, and he and von Neumann had many conversations on the subject. The conversations revealed facts and produced proofs—quite a bit of the content of the course, especially toward the end, was discovered just a week or two or three before it appeared on the blackboard. The original version of these notes was prepared by Paul Halmos, von Neumann's assistant that year. Von Neumann read the handwritten version before it went to the typist, and sometimes scribbled comments on the margins. On Chapter VI, the last one, he did more than scribble—he himself wrote most of it.

The notes were typed. Two or three copies were kept in the Institute—von Neumann had one and the Institute library had another. Since then a few photocopies have been made, but until now the notes have never been published in any proper sense of the word.

Publisher's Note

This publication was made from a copy of a manuscript titled, *Invariant Measures* by John von Neumann, Notes by Paul R. Halmos. The copy, made by Roy L. Adler, was a xerographic copy of a xerographic copy of an ozalid copy (a copying process predating xerography) of a mimeograph copy of the first five chapters and a carbon copy of the sixth. The mimeograph copy and the carbon copy were supplied by Shizuo Kakutani. As a result of copying copies of copies, a certain amount of degradation had taken place, making some of the math difficult to read, and there are some errors in the original manuscript. To have a chance at catching them all one would have had to go over the galleys with an author's dedication. No one volunteered for that kind of labor. However, we thank Roy Adler, Bruce Kitchens, Karl Petersen, and Benjamin Weiss for their substantial efforts in proofreading.

We believe most introduced typos have been caught. Those found and believed to be errors by the author have been corrected.

Another problem with the original manuscript was the use of set theory notation no longer in fashion as well as notational and typographical inconsistencies. For example, set inclusion sometimes appears as $\subseteq$ and other times as $\leq$. The manuscript was prepared in the days before word processing and TEX, and symbols were inserted by hand. Different hands were at work in inscribing them. The handwriting in Chapters IV, V, and VI is not the same as that in the first three chapters. It appears to be Halmos's handwriting first and von Neumann's later. In addition, Chapter VI is quite different in its style, notation, and numbering scheme, and almost independent from the previous chapters. Since the author could not give us his wishes regarding changes, we have kept the old fashioned usage and inconsistencies to preserve a sense of history, and in the belief that no confusion will result. We have included examples of six pages, one from each chapter, from the fourth-generation manuscript used to create this volume: page 7 from Chapter I; page 39 from Chapter II; page 56 from Chapter III; page 81 from Chapter IV; page 136 from Chapter V; and page 151 from Chapter VI. The numbers inscribed on the last chapter were not put there by von Neumann, however. Originally, the carbons were unnumbered so Adler had to number them to prevent disarray.

§3. Measurability

(3.1.1) A set M is measurable if for every set K we have

$$\nu(K) = \nu(KM) + \nu(K\tilde{M}).$$

(3.1.2) In addition to measurable sets it will also be convenient to consider measurable partitions. A partition is a finite or countable sequence of pairwise disjoint sets whose sum is S. If $\mathfrak{A}$ is the partition $(A_1, A_2, \ldots)$ and $\mathfrak{B} = (B_1, B_2, \ldots)$, we write $\mathfrak{A} \leqq \mathfrak{B}$ if every A is a subset of some B. Under this partial ordering the set of all partitions is a lattice: i.e. to every pair $\mathfrak{A}$, $\mathfrak{B}$ of partitions there corresponds a unique partition $\mathfrak{C}$ (called the product of $\mathfrak{A}$ and $\mathfrak{B}$, $\mathfrak{C} = \mathfrak{A} \cdot \mathfrak{B}$) with the properties that $\mathfrak{C} \leqq \mathfrak{A}$, $\mathfrak{C} \leqq \mathfrak{B}$, and $\mathfrak{C}' \leqq \mathfrak{A}$, $\mathfrak{C}' \leqq \mathfrak{B}$ implies $\mathfrak{C}' \leqq \mathfrak{C}$. $\mathfrak{C}$ is the partition whose sets are A_iB_j, $i,j = 1, 2, \ldots$. A partition $\mathfrak{A} = (A_1, A_2, \ldots)$ is measurable if for every set K we have

$$\nu(K) = \nu(KA_1) + \nu(KA_2) + \ldots .$$

We observe that M is a measurable set if and only if the partition $(M, \tilde{M})$ is a measurable partition.

(3.2) If M is such that $\nu(O) \geqq \nu(OM) + \nu(O\tilde{M})$ for every open set O, then M is measurable.

Proof. Let K be an arbitrary set, and O an open set, $O \supseteqq K$. Then

$$\nu(O) \geqq \nu(OM) + \nu(O\tilde{M}) \geqq \nu(KM) + \nu(K\tilde{M}).$$

Since $\nu(O) = \mu(O)$, we have

$$\mu(O) \geqq \nu(KM) + \nu(K\tilde{M})$$

for all $O \supseteqq K$, so that

$$\nu(K) = \inf \mu(O) \geqq \nu(KM) + \nu(K\tilde{M}).$$

The opposite inequality follows from (2.8.1).

FIGURE 1. Chapter I, page 7

Proof. Let $\mathfrak{B}$ be an arbitrary family of closed sets $F \subseteq S$, such that for $F_1, \ldots, F_n \in \mathfrak{B}$ always $F_1 \cdot \ldots \cdot F_n \neq \Theta$. By adding all these sets $F_1 \cdot \ldots \cdot F_n$ to $\mathfrak{B}$ we see that there is no loss of generality in assuming that $F, G \in \mathfrak{B}$ imply $F \cdot G \in \mathfrak{B}$. And still $\Theta \notin \mathfrak{B}$.

For each $F \in \mathfrak{B}$ select an element x_F^o of F.

Consider a family $\mathfrak{B}' \subseteq \mathfrak{B}$ with this property:

(10.7.1.3) There exists an $F_o \in \mathfrak{B}$ such that $F \in \mathfrak{B}'$ implies $x_F^o \notin F_o$.

The set $\mathfrak{T}$ of all such $\mathfrak{B}'$ is an ideal of subsets of $\mathfrak{B}$: That $\mathfrak{B}' \in \mathfrak{T}$ and $\mathfrak{B}'' \subseteq \mathfrak{B}'$ imply $\mathfrak{B}'' \in \mathfrak{T}$ is clear. And if $\mathfrak{B}', \mathfrak{B}'' \in \mathfrak{T}$ then $\mathfrak{B}' + \mathfrak{B}'' \in \mathfrak{T}$ because if (10.7.1.3) holds for $\mathfrak{B}'$ with F_o', and for $\mathfrak{B}''$ with F_o'', then it holds for $\mathfrak{B}' + \mathfrak{B}''$ with $F_o' \cdot F_o''$. Furthermore $x_F^o \in F$ excludes that (10.7.1.3) be true for $\mathfrak{B}$ with any F_o, hence $\mathfrak{B} \notin \mathfrak{T}$, i.e. $\mathfrak{T} \neq \mathfrak{l}$.

So we may apply our hypothesis to $I = \mathfrak{B}$ and this $\mathfrak{T}$ and obtain a function $\varphi(\xi) = \varphi(x_F | F \in \mathfrak{B})$ which fulfills (10.7.1.1), (10.7.1.2). Since $x_F^o \in F \subseteq C$, we can form $\varphi(x_F^o | F \in \mathfrak{B})$.

Consider an $F_o \in \mathfrak{B}$. Let $\mathfrak{B}'$ be the set of all $F \in \mathfrak{B}$ with $x_F^o \notin F_o$. Then $\mathfrak{B}' \in \mathfrak{T}$. Choose a $G_o \notin \mathfrak{B}'$ and form $\xi' = (x_F' | F \in \mathfrak{B})$ with

$$x_F' = x_F^o \text{ for } F \notin \mathfrak{B}',$$
$$x_F' = x_{G_o}^o \text{ for } F \in \mathfrak{B}'.$$

Then every x_F' is an x_G^o with $G \notin \mathfrak{B}'$, i.e. with $x_G^o \in F_o$, so

FIGURE 2. Chapter II, page 39

53 75
63 77
64
67
74

Chapter III

Haar measure

§13. Remarks on measures

We return to the considerations and notations of §2. We assume that S is a Hausdorff space, $\lambda(C)$ is a set function defined for all compact sets C, and $\nu(M)$ is the measure generated by $\lambda(C)$ (cf. 4.2). For convenience of reference we give below a set of properties of the space S, and the functions $\lambda(C)$ and $\nu(M)$, and then we establish certain implication relations among them.

(13.1.1) $0 \leqq \lambda(C) \leqq \infty$

(13.1.2) $\lambda(C + D) \leqq \lambda(C) + \lambda(D)$.

(13.1.3) If $CD = \Theta$, $\lambda(C + D) = \lambda(C) + \lambda(D)$.

(13.1.4) If $C \leqq D$, $\lambda(C) \leqq \lambda(D)$.

(13.1.5) If $C^i \neq \Theta$, $\lambda(C) > 0$.

(13.1.6) $\lambda(C) < \infty$.

(13.1.7) $x \to \varphi(x)$ is a homeomorphism of S into itself for which $\lambda(\varphi(C)) = \lambda(C)$ for all C.

(13.2) S is locally compact.

(13.2.1) $0 \leqq \nu(M) \leqq \infty$.

(13.3.2) $\nu(\sum_{i=1}^{\infty} M_i) \leqq \sum_{i=1}^{\infty} \nu(M_i)$.

(13.3.3) If $\{M_j\}$ is a sequence of measurable sets (cf. (13.1.1)) such that for $k \neq j$, $M_k \cdot M_j = \Theta$, then $\nu(\sum_{j=1}^{\infty} M_j) = \sum_{j=1}^{\infty} \nu(M_j)$.

(13.3.4) If $M^i \neq \Theta$, $\nu(M_i) > 0$.

(13.3.5) If $\overline{M}$ is compact, $\nu(M) < \infty$.

(13.3.6) $x \to \varphi(x)$ is a homeomorphism of S into itself for which $\nu(\varphi(M)) = \nu(M)$ for all M.

FIGURE 3. Chapter III, page 56

(17.7) <u>If $\mathcal{U}$ is a class of type $\mathcal{D}$ then $\mathcal{F}(\mathcal{U}) = \mathcal{D}'(\mathcal{U})$</u>.

<u>Proof</u>. It is clear that $\mathcal{D} = \mathcal{D}'(\mathcal{U}) \subseteq \mathcal{F} = \mathcal{F}(\mathcal{U})$; we shall prove that $\mathcal{D} = \mathcal{F}$ by showing that $\mathcal{D}$ is a field.

(17.7.1) It is clear from the definition of $\mathcal{D}(\mathcal{U})$ that $D_i \in \mathcal{D}$ for $i = 1,\ldots,n$, $D_iD_j = \theta$ for $i \neq j$, implies $D_1 + \ldots + D_n \in \mathcal{D}$

(17.7.2) If $A, B \in \mathcal{U}$ then $A\tilde{B} \in \mathcal{D}$

(17.7.3) If $A \in \mathcal{D}$ and $B \in \mathcal{U}$ then $A\tilde{B} \in \mathcal{D}$ For by hypothesis we may write A as a disjoint sum of sets of $\mathcal{U}$ $A = A_1 + A_2 + \ldots + A_n$, so that

$$A\tilde{B} = (A_1 + \ldots + A_n)\tilde{B} = A_1\tilde{B} + \ldots + A_n\tilde{B}.$$

By (17.7.2), $A_i\tilde{B} \in \mathcal{D}$ and it follows from (17.7.1) that $A\tilde{B} \in \mathcal{D}$

(17.7.4) If $A, B \in \mathcal{D}$ then $A\tilde{B} \in \mathcal{D}$. For if $B = B_1 + \ldots + B_n$, where the B_i are pairwise disjoint sets of $\mathcal{U}$, then

$$A\tilde{B} = A(\widetilde{B_1 + \ldots + B_m}) = A\tilde{B}_1\tilde{B}_2 \ldots \tilde{B}_m,$$

and the desired result follows by repeated application of (17.7.3).

(17.7.5) If $A, B \in \mathcal{D}$ $A + B \in \mathcal{D}$. For we have $A + B = A\tilde{B} + B$. The lat sum has disjoint addends which (by (17.7.4)) belong to $\mathcal{D}$ hence (by (17.7.1)) it belongs to $\mathcal{D}$

Together the statements (17.7.4) and (17.7.5) merely assert that is a field, as was to be proved.

(17.8) <u>If $\mathcal{F}$ is a field, then $\mathcal{B}(\mathcal{F}) = \mathcal{M}(\mathcal{F})$</u>

<u>Proof</u>. The structure of this proof is similar to the one given above. We observe that $\mathcal{M} = \mathcal{M}(\mathcal{F}) \subseteq \mathcal{B} = \mathcal{B}(\mathcal{F})$ and we shall complete the proof by showing that $\mathcal{M}$ is a Borel field. We remark that it is sufficient to prove that $\mathcal{M}$ is a field. For if $\mathcal{M}$ is a field and $A_i \in \mathcal{M}$, $i = 1, 2, \ldots$, then $A_i' = A_1 + \ldots + A_i \in \mathcal{M}$, whence (since $\mathcal{M}$

FIGURE 4. Chapter IV, page 81

(25.1.1) Given two $f, g \in \mathcal{H}$ and an $\epsilon > 0$, denote by $N_1{}^c(f, g; \epsilon)$ the set of all $b \in \mathcal{J}$ with

$$|(U_b f, g)| \geq \epsilon.$$

Then A has a compact closure if and only if some $N_1{}^c(f, g; \epsilon) \geqq A$.

(25.1.2) The same is true if we restrict ourselves to the $N_1{}^c(f, f; \epsilon)$ with $f = g$. We can also assume that $|f| = 1$.

(25.1.3) For $|f| = 1$ the above $N_1{}^c(f, f; \epsilon)$ is the set of all $b \in \mathcal{J}$ with

(25.1.3.1) $$|(U_b f, f)| \geq \epsilon$$ We can replace it by the set $N_2{}^c(f, \epsilon)$ of all $b \in \mathcal{J}$ with (25.1.3.2) $\mathcal{B}(U_b f, f) \geq \epsilon$.

(25.1.4) Given two Borel sets $M, N \in \mathcal{H}$ and an $\epsilon > 0$, denote by $N_3{}^c(M, N; \epsilon)$ the set of all $b \in \mathcal{J}$ with

$$\mathcal{N}(bM \cdot N) \geq \epsilon$$

Then A has a compact closure if and only if some $N_3{}^c(M, N; \epsilon) \geqq A$.

(25.1.5) The same is true if we restrict ourselves to the $N_3{}^c(M, N; \epsilon)$ with $M = N$.

Proof: We must prove two things:

(α) Each one of the above sets N^c has a compact closure.

(β) If C is compact, then there exists a set $N^c \geqq C$, for each one of the above described categories of sets N

Proof of (α): In this case it suffices to prove (25.1.1), the others are special cases of this. Indeed: (25.1.2) is a special case of (25.1.1). In (25.1.3), (25.1.3.1) is an obvious restatement of the definition of $N_1{}^c(f, f; \epsilon)$. The $N_2{}^c(f; \epsilon)$ of (25.1.3.2) may be used since (25.1.3.2) implies (25.1.3.1) so that $N_1{}^c(f, f; \epsilon) \geqq N_2{}^c(f; \epsilon)$. (25.1.4) is a special case of (25.1.1), with $f = \chi_M$, $g = \chi_N$. (25.1.5) is a special case of (25.1.4).

FIGURE 5. Chapter V, page 136

(30) (28) implies $j = k''$ for $k = 1, 2, \cdots, \tilde{p}+1$. Thus $\alpha) - \gamma)$ are satisfied (by (26), (29), (30)) with $p = \tilde{p}+1$. This contradicts our original assumption.

Thus all alternatives and Rules are exhausted, and the proof is completed.

3. Notations (Topology and Group Theory)

G : Topological group.

xy : Composition rule (in G).

x^{-1} : Reciprocal (in G).

1 Unit (in G).

M, N Arbitrary subset of G.

O, P, Q : Open subset of G.

C, D, E : Compact subsets of G.

$\overline{M}$; Closure of M (in G).

M^i ; Interior of M (in G).

xM : Set $(xu \mid u \in M)$.

Mx : Set $(ux \mid u \in M)$.

M^{-1} : Set $(u^{-1} \mid u \in M)$.

Hypotheses:

1) x^{-1} is a continuous (1-variable) function of x (in all G).

2) xy is a continuous (2-variable) function of x, y (in all G).

3) G is locally compact; i.e., there exists a C with $1 \in C^i$.

4. Equidistribution

Let a C be given which will remain fixed throughout all our discussions.

We define:

FIGURE 6. Chapter VI, p. 151

CHAPTER I

Measure Theory

1. Topology

Throughout this chapter we shall work with a fixed Hausdorff space S. The purpose of this first section is to establish the few topological lemmas needed in our considerations of measures.

1.1. A subset $C \subseteqq S$ is *compact* if, whenever $\mathfrak{A}$ is a family of spaces covering C, $C \subseteqq \sum_{\sigma\in\mathfrak{A}} \sigma$, then there exist a finite number $\sigma_1, \ldots, \sigma_n$ of open sets in $\mathfrak{A}$ covering C, $C \subseteqq \sigma_1 + \cdots + \sigma_n$. S is *locally compact* if every point of S has a neighborhood whose closure is compact.

1.2. We establish the following conventions of notation. The letters C, D, E will always denote compact sets O, P, Q will be open sets, M, N, K will be arbitrary sets. If M is any set, we denote by $\overline{M}$ the closure of M, by $\widetilde{M}$ the complementary set $S - M$, and by M^i the largest open set contained in M—the interior of M. We use the symbol θ for the empty set.

1.3. *If E is a compact set, M an arbitrary set, $E \cdot M = \theta$, and if to every point $x \in E$ there correspond disjoint open sets Q_x and R_x, $x \in Q_x$, $M \subseteqq R_x$, then there exists a pair of disjoint open sets Q and R such that $E \subseteqq Q, M \subseteqq R$.*

PROOF. Since $E \subseteqq \sum_{x\in E} Q_x$, and E is compact, we may find a finite number of points $x_1, \ldots, x_n$ such that $E \subseteqq Q_{x_1} + \cdots + Q_{x_n}$. We write $Q = Q_{x_1} + \cdots + Q_{x_n}$, $R = R_{x_1} \cdot \cdots \cdot R_{x_n}$, and verify that Q and R have the desired properties.

This lemma asserts, in less formal language, that if every point of a compact set can be separated from some fixed set M, then the whole compact set can be separated from M.

1.4. *If C and D are disjoint compact sets there exist disjoint open sets O and P such that $C \subseteqq O, D \subseteqq P$.*

PROOF. Applying 1.3 with an arbitrary point $x \in C$ in place of M, and D in place of E, we find that x can be separated from D. Since this is true for all $x \in C$ we may apply 1.3 with C in place of E and D in place of M, to obtain the desired conclusion.

1.5. *If C is a compact set and F a closed set, CF is compact.*

PROOF. Let $\mathfrak{A}$ be a family of open sets covering CF,

$$CF \subseteqq \sum_{\sigma\in\mathfrak{A}} \sigma.$$

Then we have

$$C = CF + C\widetilde{F} \subseteqq \sum_{\sigma\in\mathfrak{A}} O + \widetilde{F},$$

so that C is covered by the family $\mathfrak{A}$ and the open set $\widetilde{F}$. Since C is compact we may find $O_1, \ldots, O_n$ in $\mathfrak{A}$ such that

$$C \subseteqq O_1 + \cdots + O_n + \widetilde{F},$$

whence

$$CF \subseteqq O_1 + \cdots + O_n,$$

as was to be proved.

1.6. *If $C \subseteqq O + P$ (where C is compact and O, P are open), we may find compact sets $D \subseteqq O$, $E \subseteqq P$, for which $C \subseteqq D + E$.*

PROOF. According to 1.5, $C\widetilde{O}$ and $C\widetilde{P}$ are compact sets; since $C \subseteqq O + P$ they are disjoint. Hence, by 1.4 we may find disjoint open sets O_1 and P_1 such that $C\widetilde{O} \subseteqq O_1$, $C\widetilde{P} \subseteqq P_1$. We write $D = C\widetilde{O}_1$, $E = C\widetilde{P}_1$. Then, since $C\widetilde{O} \subseteqq O_1$, we have $\widetilde{C} + O \supseteqq \widetilde{O}_1$, so that $O \supseteqq C\widetilde{O}_1$; similarly $P \supseteqq C\widetilde{P}_1$. Furthermore, since O_1 and P_1 are disjoint, we have

$$D + E = C(\widetilde{O}_1 + \widetilde{P}_1) = C(\widetilde{O_1 P_1}) = C.$$

1.7. *If S is locally compact, and C and O are a compact and an open set respectively with $C \subseteqq O$, then there exists a compact set D such that*

$$C \subseteqq D^i \subseteqq D \subseteqq O. \tag{1.7.1}$$

PROOF. Assume first that S is compact. Then $\widetilde{O}$ is a closed set in a compact space and is therefore (see 1.5) compact; furthermore C and $\widetilde{O}$ are disjoint. By 1.4 we may find disjoint open sets Q and R such that $C \subseteqq Q$, $\widetilde{O} \subseteqq R$. Since $Q \subseteqq \widetilde{R}$, and $\widetilde{R}$ is closed, we have $\overline{Q} \subseteqq \widetilde{R}$; since in turn $\widetilde{R} \subseteqq O$, $\overline{Q} \subseteqq O$. $\overline{Q}$ is the desired set D.

If we only know about S that it is locally compact, we proceed as follows: To every point x there corresponds one of its neighborhoods N_x such that $\overline{N}_x$ is compact. If $x \in C$, we write $Q_x = O \cdot N_x$: then Q_x is a neighborhood of x whose closure is compact (see 1.5) and which is entirely included in O. Since

$$C \subseteqq \sum_{x\in C} Q_x,$$

and since C is compact, we find a finite number of points $x_1, \ldots, x_n$ such that

$$C \subseteqq T = Q_{x_1} + \cdots + Q_{x_n} \subseteqq O.$$

Then $\overline{T} = \overline{Q}_{x_1} + \cdots + \overline{Q}_{x_n}$ is a compact set. We apply the result of the preceding paragraph to the sets $C \subseteqq T$ considered as subsets of the compact space $\overline{T}$, and obtain a compact set D for which

$$C \subseteqq D^i \subseteqq D \subseteqq T.$$

Since $T \subseteqq O$, this proves everything.

2. Measure

In this section and the next we generalize to S the construction of a Lebesgue measure as usually carried out in Euclidean spaces.

2.1. Let $\lambda(C)$ be a set function defined for all compact sets C such that

$$0 \leqq \lambda(C) \leqq +\infty, \tag{2.1.1}$$

$$\lambda(C+D) \leqq \lambda(C)+\lambda(D), \tag{2.1.2}$$

$$\lambda(C+D) = \lambda(C)+\lambda(D) \quad \text{if } C \cdot D = \theta. \tag{2.1.3}$$

By means of $\lambda(C)$ we define two more set functions $\mu(O)$, for all open sets O, and $\nu(M)$ for all sets M, as follows:

$$\mu(O) = \sup_{C \subseteqq O} \lambda(C), \tag{2.1.4}$$

$$\nu(M) = \inf_{M \in O} \mu(O). \tag{2.1.5}$$

We proceed systematically to derive the properties of $\mu(O)$ and $\nu(M)$ and to establish the relations among $\lambda(C), \mu(O), \nu(M)$.

2.2.1. *If $O \subseteqq P$, then $\mu(O) \leqq \mu(P)$.*

Proof. If $C \subseteqq O$, then $C \subseteqq P$, so that, by (2.1.4), $\lambda(C) \leqq \mu(P)$, whence $\mu(O) = \sup \lambda(C) \leqq \mu(P)$.

2.2.2. *If $M \subseteqq N$, then $\nu(M) \leqq \nu(N)$.*

Proof. If $O \supseteqq N$, then $O \supseteqq M$, so that, by (2.1.5), $\mu(O) \geqq \nu(M)$, whence $\nu(N) = \inf \mu(O) \geqq \nu(M)$.

2.3.1. $\nu(C) \geqq \lambda(C)$.

Proof. If $O \supseteqq C$, then $\mu(O) \geqq \lambda(C)$, whence $\nu(C) = \inf \mu(O) \geqq \lambda(C)$.

2.3.2. $\nu(O) = \mu(O)$.

Proof. Since $O \subseteqq O$, we have, by (2.1.5), $\nu(O) \geqq \mu(O)$. If $O \subseteqq P$, then by 2.2.1, $\mu(O) \leqq \mu(P)$ so that $\mu(O) \leqq \inf \mu(P) = \nu(O)$.

2.4. $\lambda(\theta) = \mu(\theta) = \nu(\theta) = 0$.

Proof. For λ this follows from (2.1.3) by taking $C = D = \theta$. Since θ is the only compact set contained in θ, (2.1.4) implies $\mu(\theta) = \lambda(\theta)$; and, finally, since θ is open, 2.3.2 implies $\nu(\theta) = \mu(\theta)$.

2.5.1. $\mu(O+P) \leqq \mu(O) + \mu(P)$.

Proof. Suppose $C \subseteqq O + P$. Then, by 1.6, we may find $D \subseteqq O$ and $E \subseteqq P$ such that $C \subseteqq D + E$. Then (2.1.2) implies $\lambda(C) \leqq \lambda(D) + \lambda(E)$, so $\lambda(C) \leqq \sup \lambda(D) + \sup \lambda(E) = \mu(O) + \mu(P)$, whence, finally,

$$\mu(O+P) = \sup \lambda(C) \leqq \mu(O) + \mu(P).$$

We note that this proof was the first place where we made use of the relation between our set functions and the topology.

2.5.2. *If $O \cdot P = \theta$, then $\mu(O+P) = \mu(O) + \mu(P)$.*

PROOF. Suppose $C \subseteqq O+P$; write $D = C\widetilde{P}$, $E = C\widetilde{O}$. Then, since $C \subseteqq O+P$, $D \cdot E = C \cdot \widetilde{O} \cdot \widetilde{P} = C(\widetilde{O+P}) = \theta$ and, by 1.5, D and E are compact. Furthermore, $D \subseteqq O$, $E \subseteqq P$, and, since $O \cdot P = \theta$, $C = D + E$. It follows from (2.1.3) that

$$\lambda(C) = \lambda(D) + \lambda(E)$$

whence

$$\mu(O+P) = \sup \lambda(C) \geqq \lambda(D) + \lambda(E),$$

and, finally,

$$\mu(O+P) \geqq \sup \lambda(D) + \sup \lambda(E) = \mu(O) + \mu(P).$$

The desired result now follows from 2.5.1.

2.6.1.

$$\mu(O_1 + \cdots + O_n) \leqq \mu(O_1) + \cdots + \mu(O_n).$$

2.6.2. *If* $O_i \cdot O_j = \theta$ *for* $i \neq j$, $i, j = 1, \ldots, n$, *then* $\mu(O_1 + \cdots + O_n) = \mu(O_1) + \cdots + \mu(O_n)$.

The proofs of 2.6.1 and 2.6.2 are immediate by induction from 2.5.1 and 2.5.2.

2.7.1. $\mu(\sum_{i=1}^{\infty} O_i) \leqq \sum_{i=1}^{\infty} \mu(O_i)$.

PROOF. Let C be a compact set, $C \subseteqq \sum_{i=1}^{\infty} O_i$; then we can find a finite number of subscripts $i_1, \ldots, i_n$ such that

$$C \subseteqq O_{i_1} + \cdots + O_{i_n}.$$

It follows from (2.1.4) that

$$\lambda(C) \leqq \mu(O_{i_1} + \cdots + O_{i_n}) \leqq \mu(O_{i_1}) + \cdots + \mu(O_{i_n}),$$

so that

$$\lambda(C) \leqq \sum_{i=1}^{\infty} \mu(O_i).$$

Since this is true for all $C \subseteqq \sum_{i=1}^{\infty} O_i$, we have

$$\mu\left(\sum_{i=1}^{\infty} O_i\right) = \sup \lambda(C) \leqq \sum_{i=1}^{\infty} \mu(O_i).$$

2.7.2. *If* $O_i \cdot O_j = \theta$ *for* $i \neq j$, $i, j = 1, 2, \ldots$, *then* $\mu(\sum_{i=1}^{\infty} O_i) = \sum_{i=1}^{\infty} \mu(O_i)$.

PROOF. We have

$$\mu\left(\sum_{i=1}^{\infty} O_i\right) \geqq \mu\left(\sum_{i=1}^{n} O_i\right) = \sum_{i=1}^{n} \mu(O_i),$$

whence, since this is true for all n,

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

The result now follows from 2.7.1.

2.8.1. $\nu(M_1 + M_2 + \cdots) \leqq \nu(M_1) + \nu(M_2) + \cdots$.

PROOF. For any $\varepsilon > 0$ we may find open sets $O_i \supseteqq M_i$ such that $\mu(O_i) \leqq \nu(M_i) + \varepsilon/2^i$, for $i = 1, 2, \ldots$. Then we have

$$\nu\left(\sum_{i=1}^{\infty} M_i\right) \leqq \mu\left(\sum_{i=1}^{\infty} O_i\right) \leqq \sum_{i=1}^{\infty} \mu(O_i)$$

$$\leqq \sum_{i=1}^{\infty} \nu(M_i) + \varepsilon.$$

Since ε was arbitrary, the result follows.

3. Measurability

3.1.1. A set M is *measurable* if for every set K we have

$$\nu(K) = \nu(KM) + \nu(K\widetilde{M}).$$

3.1.2. In addition to measurable sets it will also be convenient to consider measurable partitions. A *partition* is a finite or countable sequence of pairwise disjoint sets whose sum is S. If $\mathfrak{A}$ is the partition $(A_1, A_2, \ldots)$ and $\mathfrak{B} = (B_1, B_2, \ldots)$, we write $\mathfrak{A} \leqq \mathfrak{B}$ if every A is a subset of some B. Under this partial ordering the set of all partitions is a *lattice*: i.e. to every pair $\mathfrak{A}, \mathfrak{B}$ of partitions there corresponds a unique partition $\mathfrak{C}$ (called the product of $\mathfrak{A}$ and $\mathfrak{B}$, $\mathfrak{C} = \mathfrak{A} \cdot \mathfrak{B}$) with the properties that $\mathfrak{C} \leqq \mathfrak{A}$, $\mathfrak{C} \leqq \mathfrak{B}$, and $\mathfrak{C}' \leqq \mathfrak{A}$, $\mathfrak{C}' \leqq \mathfrak{B}$ implies $\mathfrak{C}' \leqq \mathfrak{C}$. $\mathfrak{C}$ is the partition whose sets are A_iB_j, $i, j = 1, 2, \ldots$. A partition $\mathfrak{A} = (A_1, A_2, \ldots)$ is *measurable* if for every set K we have

$$\nu(K) = \nu(KA_1) + \nu(KA_2) + \cdots.$$

We observe that M is a measurable set if and only if the partition $(M, \widetilde{M})$ is a measurable partition.

3.2. *If M is such that $\nu(O) \geqq \nu(OM) + \nu(O\widetilde{M})$ for every open set O, then M is measurable.*

PROOF. Let K be an arbitrary set, and O an open set, $O \supseteqq K$. Then

$$\nu(O) \geqq \nu(OM) + \nu(O\widetilde{M}) \geqq \nu(KM) + \nu(K\widetilde{M}).$$

Since $\nu(O) = \mu(O)$, we have

$$\mu(O) \geqq \nu(KM) + \nu(K\widetilde{M})$$

for all $O \supseteqq K$, so that

$$\nu(K) = \inf \mu(O) \geqq \nu(KM) + \nu(K\widetilde{M}).$$

The opposite inequality follows from 2.8.1.

3.3. *The product of two measurable partitions is measurable.*

PROOF. If $\mathfrak{A} = (A_1, A_2, \ldots)$ and $\mathfrak{B} = (B_1, B_2, \ldots)$ are measurable, and M is an arbitrary set, we have

$$\nu(M) = \sum_{i=1}^{\infty} \nu(MA_i) = \sum_{i=1}^{\infty}\sum_{j=1}^{\infty} \nu(MA_iB_j).$$

It follows from 3.1.2 that $\mathfrak{A} \cdot \mathfrak{B}$ is measurable.

3.4. *If $\mathfrak{A}$ is a measurable partition and $\mathfrak{A} \leqq \mathfrak{B}$, then $\mathfrak{B}$ is measurable.*

PROOF. For any set M we have

$$\nu(M) = \sum_{i=1}^{\infty} \nu(MA_i) = \sum_{j=1}^{\infty} \sum_{A_i \subseteqq B_j} \nu(MA_i).$$

On the other hand,

$$\nu(MB_j) = \sum_{i=1}^{\infty} \nu(MA_iB_j) = \sum_{A_i \subseteqq B_j} \nu(MA_iB_j) = \sum_{A_i \subseteqq B_j} \nu(MA_i),$$

so that

$$\nu(M) = \sum_{j=1}^{\infty} \nu(MB_j),$$

as was to be proved.

3.5. *If $\mathfrak{A} = (A_1, A_2, \ldots)$, a necessary and sufficient condition for the measurability of $\mathfrak{A}$ is that of every A_i.*

PROOF. If $\mathfrak{A}$ is measurable, so is the partition $(A_i, \widetilde{A}_i)$ for every A_i (by 3.4), and therefore so is A_i. Conversely, suppose every A_i is measurable. Then each of the partitions $(A_1, \widetilde{A}_1), \ldots, (A_n, \widetilde{A}_n)$ is measurable and therefore, by 3.3, so is their product. Since the A_j are pairwise disjoint, this product is the partition

$$\left(A_1, A_2, \ldots, A_n, \prod_{j=1}^{n} \widetilde{A}_j\right).$$

Hence, for every set M,

$$\nu(M) = \nu(MA_1) + \cdots + \nu(MA_n) + \nu\left(M \cdot \prod_{j=1}^{n} \widetilde{A}_j\right) \geqq \sum_{i=1}^{n} \nu(MA_i),$$

so that $\nu(M) \geqq \sum_{i=1}^{\infty} \nu(MA_i)$. Since the opposite inequality is true by 2.8.1, $\mathfrak{A}$ is measurable.

The three results 3.3, 3.4, and 3.5 can now be used to derive easily all the essential properties of measurable sets and the behavior of 2 on the measurable sets.

3.6. *If A and B are measurable sets so are $\widetilde{A}, AB, A + B$.*

PROOF. For $\widetilde{A}$ this follows directly from the symmetry in A and $\widetilde{A}$ of the definition 3.1.1. Since by hypothesis $(A, \widetilde{A})$ and $(B, \widetilde{B})$ are measurable partitions and AB is one of the sets in their product, AB is measurable (by 3.4). The measurability of $A + B$ is now obvious.

3.7. *If $A_1, A_2, \ldots$ are measurable sets and $A = \sum_{i=1}^{\infty} A_i$, then A is measurable; if the A_n are pairwise disjoint, then*

$$\nu(A) = \nu(A_1) + \nu(A_2) + \cdots. \tag{3.7.1}$$

PROOF. Suppose first that A_i are disjoint. Since A_i is measurable, the partition $(A_i, \widetilde{A}_i)$ is measurable, and therefore the product

$$\left(A_1, A_2, \ldots, A_n, \prod_{j=1}^{n} \widetilde{A}_j\right)$$

of the first n of these partitions is measurable. Hence, if M is any set,

$$\nu(M) = \sum_{j=1}^{n} \nu(MA_j) + \nu\left(M \cdot \prod_{j=1}^{n} \widetilde{A}_j\right) \geqq \sum_{j=1}^{n} \nu(MA_j) + \nu\left(M \cdot \prod_{j=1}^{\infty} \widetilde{A}_j\right),$$

and since this is true for every n,

$$\nu(M) \geqq \sum_{j=1}^{\infty} \nu(MA_j) + \nu\left(M \cdot \prod_{j=1}^{\infty} \widetilde{A}_j\right).$$

The opposite inequality holds by 2.3.1, so that we have

$$\nu(M) = \sum_{j=1}^{\infty} \nu(MA_j) + \nu\left(M \cdot \prod_{j=1}^{\infty} \widetilde{A}_j\right). \tag{3.7.2}$$

This means that the partition

$$\left(A_1, A_2, \ldots, \prod_{j=1}^{\infty} \widetilde{A}_j\right)$$

is measurable, whence, by 3.5, so is the set $\prod_{j=1}^{\infty} \widetilde{A}_j$, and therefore $\sum_{j=1}^{\infty} A_j = A$. Writing A for M in (3.7.2) yields (3.7.1).

In case the A_i are not disjoint, we observe that we may write

$$A = A_1 + \widetilde{A}_1 A_2 + \widetilde{A}_1 \widetilde{A}_2 A_3 + \cdots$$

and each term in this sum is measurable (by 3.6) and the terms are pairwise disjoint. This completes the proof of 3.7.

The complete additivity of ν for measurable sets implies (since the empty set is measurable!) its finite additivity. In the finite case, however, we can improve 3.7 as follows.

3.8. *For any two sets A and B of which one, say A, is measurable, we have*

$$\nu(A+B) + \nu(AB) = \nu(A) + \nu(B).$$

PROOF. The measurability of A implies the relations

$$\nu(A+B) = \nu((A+B)A) + \nu((A+B)\widetilde{A}) = \nu(A) + \nu(B\widetilde{A}), \tag{3.8.1}$$

$$\nu(B) = \nu(BA) + \nu(B\widetilde{A}). \tag{3.8.2}$$

It follows that

$$\nu(A) + \nu(B) = \nu(A) + \nu(BA) + \nu(B\widetilde{A}) = \nu(A+B) + \nu(AB).$$

3.9.1. *If $A_1 \subseteqq A_2 \subseteqq \cdots$ is a monotone increasing sequence of measurable sets, then*

$$\nu\left(\sum_{j=1}^{\infty} A_j\right) = \lim_{j\to\infty} \nu(A_j).$$

PROOF. Since

$$\sum_{j=1}^{\infty} A_j = A_1 + \widetilde{A}_1 A_2 + \widetilde{A}_2 A_3 + \widetilde{A}_3 A_4 + \cdots,$$

we have

$$\nu\left(\sum_{j=1}^{\infty} A_j\right) = \lim_{j\to\infty}\left(\nu(A_1) + \sum_{i=1}^{j} \nu(\widetilde{A}_i A_{i+1})\right) = \lim_{j\to\infty} \nu(A_j).$$

3.9.2. *If $A_1 \supseteqq A_2 \supseteqq \cdots$ is a monotone decreasing sequence of measurable sets of which not all have infinite measure, then*

$$\nu\left(\prod_{j=1}^{\infty} A_j\right) = \lim_{j\to\infty} \nu(A_j).$$

PROOF. If $\nu(A_n) < \infty$, then since $\nu(A_n) \geqq \nu(A_{n+p})$ for $p = 1, 2, \ldots, \nu(A_{n+p}) < \infty$. Hence there is no loss of generality in assuming that every A_n has finite measure, or, in other words, that all A_n are contained in a measurable set M of finite measure. The theorem now follows by applying 3.9.1 to the sequence $M - A_1, M - A_2, \ldots$.

We conclude this section by proving that its considerations have not been entirely vacuous.

3.10. *Every open set is measurable.*

PROOF. Let O be an open set, and choose another arbitrary open set P. Let C and D be compact sets, $C \subseteqq OP$, $D \subseteqq P\widetilde{C}$. Then $D \cdot C = \theta$ (since $D \subseteqq \widetilde{C}$), and since $C \subseteqq P$, $D \subseteqq P$, we have $C + D \subseteqq P$. Therefore

$$\nu(P) = \mu(P) \geqq \lambda(C + D) = \lambda(C) + \lambda(D),$$

so that, since $P\widetilde{C}$ is open and $\widetilde{C} \supseteqq \widetilde{O}$, we have

$$\begin{aligned}\nu(P) &\geqq \lambda(C) + \sup \lambda(D) = \lambda(C) + \mu(P\widetilde{C}) \\ &= \lambda(C) + \nu(P\widetilde{C}) \geqq \lambda(C) + \nu(P\widetilde{O}),\end{aligned}$$

and hence

$$\begin{aligned}\nu(P) &\geqq \sup \lambda(C) + \nu(P\widetilde{O}) = \mu(PO) + \nu(P\widetilde{O}) \\ &= \mu(PO) + \nu(P\widetilde{O}).\end{aligned}$$

According to 3.2 then, O is measurable.

3.11. *Every Borel set is measurable.*

PROOF. Immediate by 3.10 and 3.6, 3.7.

4. Connection between λ and ν

In 2.3.1 we saw that for compact sets C, $\nu(C) \geqq \lambda(C)$. In this section we shall establish an inequality between $\nu(C)$ and $\lambda(C)$ in the opposite direction and then discuss, under suitable hypotheses, to what extent λ is determined by ν. We recall the notation established in 1.2: for every set M we write M^i for the set of all interior points of M.

4.1. *If in addition to* (2.1.1)–(2.1.3) $\lambda(C)$ *also satisfies the condition*

(4.1.1) $$\lambda(C) \leqq \lambda(D) \quad \textit{for } C \subseteqq D,$$

then for every compact set C

(4.1.2) $$\nu(C^i) \leqq \lambda(C).$$

PROOF. According to 2.3.2 and (2.1.4) we have

$$\nu(C^i) = \mu(C^i) = \sup \lambda(D) \quad \text{for } D \subseteqq C^i,$$

since if $D \subseteqq C^i$, then $D \subseteqq C$, $\lambda(D) \leqq \lambda(C)$, whence

$$\sup \lambda(D) = \nu(C^i) \leqq \lambda(C).$$

4.2. To have a convenient terminology we shall say that a set function $\lambda(C)$ satisfying (2.1.1)–(2.1.3) *generates* $\nu(M)$ if (2.1.4) and (2.1.5) hold. We observe that to say that a measure $\nu(M)$ is the measure generated by some set function $\lambda(C)$ satisfying (2.1.1)–(2.1.3) is equivalent, in the Carathéodory terminology, to saying that $\nu(M)$ is a *regular* measure.

The object of the remainder of this section is to determine, under suitable hypotheses on the space S, which $\lambda(C)$ can generate $\nu(M)$. The first theorem below, 4.3, shows that the largest possible $\lambda(C)$ always does generate $\nu(M)$.

4.3. *If* $\lambda_1(C)$ *and* $\lambda_2(C)$ *are set functions satisfying* (2.1.1)–(2.1.3) *such that* $\lambda_1(C)$ *generates* $\nu(M)$ *and* $\lambda_1(C) \leqq \lambda_2(C) \leqq \nu(C)$ *for all compact sets* C, *then* $\lambda_2(C)$ *generates* $\nu(M)$.

PROOF. We are to prove that, for every open set O,

$$\nu(O) = \sup \lambda_2(C) \quad \text{for } C \subseteqq O.$$

We have, for every compact $C \subseteqq O$,

$$\nu(O) \geqq \nu(C) \geqq \lambda_2(C) \geqq \lambda_1(C)$$

(from 2.3.1), so that

$$\nu(O) \geqq \sup \lambda_2(C) \geqq \sup \lambda_1(C) = \nu(O),$$

and this implies the desired result.

The result proved above implies easily that if $\nu(M)$ is a measure generated by some $\lambda_0(C)$, then $\nu(M)$ is also generated by $\lambda(C)$, if $\lambda(C)$ is defined by $\lambda(C) \equiv \nu(C)$. In the following theorem we shall prove that even the smallest $\lambda(C)$ (and therefore, by 4.3, all $\lambda(C)$ in the range allowed by the inequalities 2.3.1 and (4.1.2)) will generate $\nu(M)$. We observe that we shall not need any assumption of the type of (4.1.1) restricting the $\lambda(C)$ which originally generated $\nu(M)$, but that the lower bound given by (4.1.2) is a lower bound only for those $\lambda(C)$ which satisfy (4.1.1). We shall not discuss the question as to whether or not we may find a $\lambda(C)$

satisfying (2.1.1)–(2.1.3) but *not* satisfying (4.1.2) (and, therefore, automatically, not satisfying (4.1.1)) which generates $\nu(M)$.

4.4. *If S is locally compact and $\nu(M)$ is a measure generated by a set function $\lambda_0(C)$ satisfying* (2.1.1)–(2.1.3), *then $\lambda(C) \equiv \nu(C^i)$ also generates $\nu(M)$.*

PROOF. We remark that we shall not explicitly use the hypothesis that $\nu(M)$ is generated by $\lambda_0(C)$: we assume it only to be able to apply the consequence of 4.3 already pointed out in the last paragraph, namely that $\lambda'(C) \equiv \nu(C)$ generates $\nu(M)$. This means that for any open set O

$$\nu(O) = \sup \nu(C) \quad \text{for } C \subseteqq O. \tag{4.4.1}$$

To prove 4.4 it is sufficient to prove that for every open set O we have

$$\nu(O) = \sup \nu(D^i) \quad \text{for } D \subseteqq O. \tag{4.4.2}$$

Let C be any compact set, $C \subseteqq O$. Since S is locally compact we may, according to 1.7, find a compact set D such that

$$C \subseteqq D^i \subseteqq D \subseteqq O.$$

Then $\nu(C) \leqq \nu(D^i) \leqq \nu(O)$ so that

$$\sup \nu(C) \leqq \sup \nu(D^i) \leqq \nu(O). \tag{4.4.3}$$

Since the left term of (4.4.3) is, by (4.4.1), equal to $\nu(O)$, (4.4.3) implies (4.4.2).

CHAPTER II

Generalized limits

5. Topology

In this section we state the auxiliary results from topology that we shall have occasion to use in this chapter.

5.1. *Every compact set C is closed.*

5.2. *A closed set C is compact if and only if the condition* 5.2.1 (*below*) *implies* 5.2.2.

5.2.1. *$\mathfrak{B}$ is a family of closed sets $F \subseteqq C$, such that for $F_1, \ldots, F_n \in \mathfrak{B}$ always $F_1 \cdot \cdots \cdot F_n \neq 0$.*

5.2.2. *$\mathfrak{B}$ is a family of closed sets $F \subseteqq C$, such that $\prod_{F \in \mathfrak{B}} F \neq 0$.*

PROOF. For 5.1: Consider a point $p \notin C$. For every $q \in C$ we have $q \neq p$; hence disjoint neighborhoods O_q of q and P_q of p exist. $q \in O_q$; hence $C \subseteqq \sum_{q \in C} O_q$. Since C is compact this implies by 1.1 that $C \subseteqq O_{q_1} + \cdots + O_{q_n}$. Therefore C and $P_{q_1} \cdot \cdots \cdot P_{q_n}$ are disjoint. But $P_{q_1} \cdots P_{q_n}$ is an open set and contains p. Hence $p \notin \overline{C}$. In other words $p \notin C$ implies $p \notin \overline{C}$, so that C is closed.

For 5.2: If we replace S by the closed subset C and topologize C accordingly (so that "closed" sets in C are closed sets in S which are $\subseteqq C$, and "open" sets in C are the intersections of open sets in S with C), then the conditions 5.2.1, 5.2.2 are unaffected. So is the definition of compactness in 1.1—we must only replace the sets $O \in \mathfrak{A}$ (open in S) by the corresponding $O \cdot C$ (open in C).

So we lose no generality in assuming $C = S$. But then 5.2 and 1.1 are clearly equivalent. We must only pass from the open $O \in \mathfrak{A}$ to the closed $\widetilde{O}$ and thus form $\mathfrak{B}$—or vice versa.

As an immediate corollary to 5.2 we obtain:

5.3. *Let S and T be two topological spaces and $p \to f(p)$ a continuous* (*possibly many to one*) *mapping of a $C \subseteqq S$ on a set $M \subseteqq T$. Then the compactness of C implies that of its image M.*

PROOF. We use the characterization of compactness given in 5.2. Then the observation that every closed subset of M has a closed subset of C for inverse image by the (possibly one to many) mapping $f(p) \to p$, yields the desired result.

5.4. Let I be an arbitrary set (of indices) and for each $\alpha \in I$ let S_α be a Hausdorff space. We call a function x_α whose domain is I and whose value at $\alpha_0 \in I$ lies in the space S_{α_0} a *sequence*; we denote it by $(x_\alpha \,|\, \alpha \in I)$. We now define a topology on the space Ξ of all sequences. Let $\alpha_1, \ldots, \alpha_n$ be any finite number of elements of I and let $N_1, \ldots, N_n$ be arbitrary open sets in the spaces

$S_{\alpha_1}, \dots, S_{\alpha_n}$ respectively. The set of all sequences $(x_\alpha \mid \alpha \in I)$ for which $x_{\alpha_i} \in N_i$, $i = 1, \dots, n$, we shall call a *neighborhood* of any sequence contained in it. With this definition of neighborhood the sequence space becomes a Hausdorff space. We shall generally denote points of this space by $\xi = (x_\alpha \mid \alpha \in I)$; for each $\alpha_0 \in I$ we define the function $x_{\alpha_0}(\xi)$ as the function whose value at $\xi_0 = (x_{\alpha_0} \mid \alpha \in I)$ is $x^0_{\alpha_0}$.[1] In case all the spaces S_α coincide, $S_\alpha = S$, we write $C(\xi) = C(x_\alpha \mid \alpha \in I)$ for the closure (in S) of the set of points x_α, $\alpha \in I$. In case $C(\xi)$ is compact, we call ξ a *bounded* sequence.

5.5. Retaining the notation in 5.4 we state here, without proof, a fundamental theorem concerning sequence spaces, due to Tychonoff. *If S_α is compact, for every $\alpha \in I$, then the space Ξ of all sequences $(x_\alpha \mid \alpha \in I)$ is compact.*[2]

5.6. We shall have occasion to use, in addition to the sequence space Ξ defined above, a space Φ of the same type. Φ is defined similarly to Ξ: we take I to be a fixed set $\Xi^* \subseteqq \Xi$, for each $\xi \in \Xi^*$ we take an arbitrary subspace $S(\xi)$ of S (where we assume that all $S_\alpha = S$; cf. 5.4 above), and form the class Φ of all functions $\varphi(\xi)$ whose value at $\xi = \xi_0$ lies in the space $S(\xi_0)$. It will be convenient always to refer to elements of Ξ as *sequences*, and to elements of Φ as *functions*. Topology is introduced in Φ in terms of the topology already defined on S (and thereby on $S(\xi)$) just as the topology was introduced in Ξ in terms of the given topology on S_α. Tychonoff's theorem 5.4 holds, of course, for Φ: if each $S(\xi)$ is compact, so is Φ.

We return to the consideration of Ξ.

5.7. *$x_{\alpha_0}(\xi)$ is a continuous function of ξ for every $\alpha_0 \in I$.*

PROOF. Let ξ_0 be an arbitrary sequence, $\xi_0 \in \Xi$ and let N be an arbitrary neighborhood (in S_{α_0}) of $x_{\alpha_0}(\xi_0)$. The set of all sequences ξ for which $x_{\alpha_0}(\xi) \in N$ is a neighborhood M of ξ_0: for every $\xi \in M$, $x_{\alpha_0}(\xi) \in N$, so that $x_{\alpha_0}(\xi)$ is continuous.

5.8. The result of 5.7 is true for every space formed similarly to Ξ: in particular it is true for Φ. In other words: *for any fixed $\xi_0 \in \Xi^*$, $\varphi(\xi_0)$ considered as a function of φ (a function whose domain is Φ and whose range is $S(\xi_0)$) is continuous.*

5.9. We introduce one more topology on sequence spaces Ξ. For this purpose we consider only such spaces Ξ for which $S_\alpha = S$ is the same for all α. We shall define neighborhoods not of all sequences $\xi = (x_\alpha \mid \alpha \in I)$ (where $x_\alpha \in S$) but only for those sequences ξ_0 for which $x_\alpha(\xi_0)$ is independent of α: $x_\alpha(\xi_0) = x^0 \in S$. By a neighborhood of ξ_0 we shall mean the set of all sequences $\xi = (x_\alpha \mid \alpha \in I)$ (with not necessarily constant $x_\alpha(\xi)$) for which $x_\alpha \in O^0$ for every $\alpha \in I$, where O^0 is a

---

[1] This is an error in the original manuscript. It should read $\xi_0 = (x^0_\alpha | \alpha \in I)$ is $x^0_{\alpha_0}$.

[2] The original publication concerning this theorem is:

A. Tychonoff, *Ueber die topologische Erweiterung von Räumen*, Math. Annalen, vol. 102 (1930), pp. 544–561—specifically pp. 546, 548–550.

Considerably simpler proofs were given by:

J. W. Alexander and L. Zippin, *Discrete Abelian groups and their character groups*, Annals of Math., vol. 36 (1935), pp. 71–85—specifically pp. 75–76.

H. Wallman, *Lattices and topological spaces*, Annals of Math., vol. 39 (1938), pp. 112–126—specifically pp. 123–124.

J. W. Alexander, *Lectures on topology and lattice theory*, Princeton, 1940–41.

neighborhood of x^0 (in S). A function $f(\xi) = f(x_\alpha \mid \alpha \in I)$ will be said to be *feebly continuous* if it is continuous (in the sense of the topology just defined) at every point ξ_0 at which neighborhoods are defined.

(This terminology is at variance with the usual one: the topology defined in 5.4, which we propose to use without any adjective, is usually called weak topology, while the above one, which we termed feeble, is a generalization and an analogue of one usually called uniform topology. Our present terminology is more suggestive in our applications, where the topologies themselves are less in the foreground than the notions of continuity which they define.)

6. Ideals

Throughout this section we shall consider an arbitrary fixed set I.

6.1. A nonempty collection $\mathcal{T}$ of subsets of I is an *ideal* if with every set it contains all its subsets and with any two sets it contains their sum. $\mathcal{T}$ is a *dual ideal* if with every set it contains all its supersets and with any two sets it contains their intersection. Every ideal contains θ; every dual ideal contains I. The set of all subsets of I is both an ideal and a dual ideal. It is obviously the only ideal which contains I, and the only dual which contains θ—under both aspects we call it the *unit* and denote it by $\mathbf{I}$.

6.2. *The following five conditions for an ideal $\mathcal{T}$ are all equivalent.*

6.2.1. $\mathcal{T} \neq \mathbf{I}$; *for any set $I' \subseteqq I$, $\mathcal{T}$ contains at least one of the sets I' and $\widetilde{I'}$.*

6.2.2. $\mathcal{T} \neq \mathbf{I}$; *if $\mathcal{T}'$ is an ideal, $\mathcal{T} \subseteqq \mathcal{T}'$, then $\mathcal{T}' = \mathcal{T}$, or $\mathcal{T}' = \mathbf{I}$.*

6.2.3. *For any set $I' \subseteqq I$, $\mathcal{T}$ contains exactly one of the sets I' and $\widetilde{I'}$.*

6.2.4. *The collection $\widetilde{\mathcal{T}}$ of all sets ($\subseteqq I$) not contained in $\mathcal{T}$ is a dual ideal.*

6.2.5. $\mathcal{T} \neq \mathbf{I}$; *if $I' \cdot I'' \in \mathcal{T}$, then either $I' \in \mathcal{T}$ or else $I'' \in \mathcal{T}$.*

PROOF. We shall prove that (cyclically, in the order given) each one of the five conditions implies the next one.

(i) 6.2.1 $\to$ 6.2.2. If $\mathcal{T}'$ is an ideal, $\mathcal{T} \subseteqq \mathcal{T}'$, and if $\mathcal{T}' \neq \mathcal{T}$, then there is a set $I' \subseteqq I$, contained in $\mathcal{T}'$ but not in $\mathcal{T}$. Then $\widetilde{I'}$ must be in $\mathcal{T}$, and therefore in $\mathcal{T}'$, so that $I' + \widetilde{I'} = I \in \mathcal{T}'$. Hence $\mathcal{T}' = \mathbf{I}$.

(ii) 6.2.2 $\to$ 6.2.3. Let I' be any subset of I. The collection of all subsets of sets of the form $I' + J$, where $J \in \mathcal{T}$, is an ideal $\mathcal{T}'$ and $\mathcal{T} \subseteqq \mathcal{T}'$. If neither I' nor $\widetilde{I'}$ are contained in $\mathcal{T}$, then we should have $I' \in \mathcal{T}'$, $\mathcal{T} \neq \mathcal{T}'$. Hence, assuming 6.2.2 it follows that $\mathcal{T}' = \mathbf{I}$. Consequently $I \in \mathcal{T}'$, $I \subseteqq I' + J$ where $J \in \mathcal{T}$, and so $\widetilde{I'} \subseteqq J$, $\widetilde{I'} \in \mathcal{T}$, contradicting our above assumption. Then $\mathcal{T}$ contains at least one of the sets I' and $\widetilde{I'}$ (this, incidentally, is the statement of 6.2.1). If it contained both we would have $I' + \widetilde{I'} = I \in \mathcal{T}$; i.e. $\mathcal{T} = \mathbf{I}$, which is not the case.

(iii) 6.2.3 $\to$ 6.2.4. Replacement of every set $I' \subseteqq I$ by its complementary set I' obviously carries an ideal into a dual ideal. But owing to 6.2.3 this process carries $\mathcal{T}$ into its complementary set $\widetilde{\mathcal{T}}$.

(iv) 6.2.4 $\to$ 6.2.5. Since $\widetilde{\mathcal{T}}$ is a dual ideal, $I \in \widetilde{\mathcal{T}}$; hence $I \notin \mathcal{T}$, $\mathcal{T} \neq \mathbf{I}$. $I', I'' \notin \mathcal{T}$ means $I', I'' \in \widetilde{\mathcal{T}}$, and therefore, since $\widetilde{\mathcal{T}}$ is a dual ideal, $I' \cdot I'' \in \widetilde{\mathcal{T}}$, i.e. $I' \cdot I'' \notin \mathcal{T}$. So $I' \cdot I'' \in \mathcal{T}$ implies $I' \in \mathcal{T}$ or $I'' \in \mathcal{T}$.

(v) 6.2.5 $\to$ 6.2.1. If $I' \subseteqq I$, then, since $\mathcal{T}$ is an ideal, $\theta = I' \cdot \widetilde{I'} \in \mathcal{T}$, so that either $I' \in \mathcal{T}$ or $\widetilde{I''} \in \mathcal{T}$.

This completes the proof of 6.2.

6.3. An ideal $\mathcal{T}$ satisfying any one of the five equivalent conditions of 6.2 is called a *prime ideal.* The algebraic analogy which motivates this name is best observed in 6.2.2. This also makes it clear that the prime ideals are the maximal ideals $\neq$ **I**. Another interesting characterization is 6.2.4 which makes the relation of prime ideals to duality clearer.[3]

7. Independence

A function of a finite number of variables (i.e. of a sequence $\xi = (x_\alpha \,|\, \alpha \in I)$ for a finite I) is obviously a constant if it is independent of every one of its variables. The same is not true for a function of an infinite number of variables (i.e. for an infinite I, cf. above), as the example of $I = (1, 2, \dots)$, $S =$ set of all real numbers x, $0 \leqq x \leqq 1$, and the function $\limsup_{\alpha \to \infty} x_\alpha$ shows. Thus for functions of infinitely many variables a closer analysis of the notion of independence is necessary.

We proceed to carry out such an analysis, using the concepts introduced above. We shall make use, in particular, of 5.4 and 6.1.

7.1. If $\varphi(\xi)$ is a function defined on a set $\Xi^* \subseteqq \Xi$ (where Ξ is the sequence space of 5.3 and $\varphi(\xi)$ is therefore an element of the function space Φ of 5.5), and if I' is any subset of the index set I, we shall say that $\varphi(\xi)$ is *independent* of I' if for any two $\xi_1, \xi_2 \in \Xi^*$ the validity of $x_\alpha(\xi_1) = x_\alpha(\xi_2)$ for $\alpha \notin I'$ implies $\varphi(\xi_1) = \varphi(\xi_2)$. Throughout what follows we assume that the set $\Xi^* \subseteqq \Xi$ has the following property:

7.1.1. *If* $\xi_1 = (x^1_\alpha \,|\, \alpha \in I)$, $\xi_2 = (x^2_\alpha \,|\, \alpha \in I)$, $\xi_0 = (x^0_\alpha \,|\, \alpha \in I)$ *are three sequences such that for every* $\alpha \in I$, $x^0_\alpha = x^1_\alpha$ *or* $x^0_\alpha = x^2_\alpha$, *then* $\xi_1, \xi_2 \in \Xi^*$ *implies* $\xi_0 \in \Xi^*$.

7.2. *The collection* $\mathcal{T}$ *of all sets* $I' \subseteqq I$ *for which* $\varphi(\xi)$ *is independent of* I' *is an ideal: the independence ideal of* φ.

PROOF. That $\mathcal{T}$ contains with any set all its subsets is clear from the definition of independence; $\mathcal{T}$ is not empty since $\theta \in \mathcal{T}$. Now suppose that I' and I'' are both in $\mathcal{T}$ and that $x_\alpha(\xi_1) = x_\alpha(\xi_2)$ for $\alpha \notin I' + I''$. We define a sequence $\xi_0 = (x^0_\alpha \,|\, \alpha \in I)$ as follows:

For $\alpha \in I'$, $x^0_\alpha = x_\alpha(\xi_1)$;

for $\alpha \in \widetilde{I'} \cdot I''$, $x^0_\alpha = x_\alpha(\xi_2)$;

for $\alpha \in \widetilde{I'} \cdot \widetilde{I''}$, $x^0_\alpha = x_\alpha(\xi_1) = x_\alpha(\xi_2)$.

Then $\xi_0 \in \Xi^*$ by 7.1.1, $x_\alpha(\xi_0) = x_\alpha(\xi_1)$ for $\alpha \notin I''$, and $x_\alpha(\xi_0) = x_\alpha(\xi_2)$ for $\alpha \notin I'$, so that $\varphi(\xi_1) = \varphi(\xi_0) = \varphi(\xi_2)$. This implies that $I' + I'' \in \mathcal{T}$ and completes the proof of 7.2.

7.3. It is important to observe that the independence ideals are subject to no other restrictions. I.e.: We assume for the sake of simplicity that $\Xi^* = \Xi$ and also, omitting a trivial case, that S possesses at least two points. I, however, is perfectly arbitrary. Under these assumptions we prove:

---

[3] The use of subset prime ideals in set theory is due to M. H. Stone, *The theory of representations for Boolean algebras*, Trans. Amer. Math. Soc. **40** (1936), 37–111—specifically pp. 74–78.

Corresponding to every ideal $\mathcal{T}$ (of subsets of I) there is at least one function $\varphi(\xi) = \varphi(x_\alpha \,|\, \alpha \in I)$ such that the independence ideal of φ is $\mathcal{T}$.

PROOF. It is sufficient to consider only the case where S contains exactly two points, x' and x''. For if the theorem is proved in this case and $\varphi_0(\xi)$ is a function whose independence ideal is $\mathcal{T}$, then the general case follows by splitting S into two disjoint subsets S' and S'' and defining $\varphi(x_\alpha \,|\, \alpha \in I) = \varphi_0(y_\alpha \,|\, \alpha \in I)$, where $y_\alpha = x'$ or x'' according as $x_\alpha \in S'$ or $x_\alpha \in S''$.

Suppose then that x_α takes only two values, say 0 and 1, for all $\alpha \in I$. Then any function $\varphi(x_\alpha \,|\, \alpha \in I)$ depends only on the set K of α's for which $x_\alpha = 1$, and conversely any set function $f(K)$ determines, by this correspondence, a unique φ. We assert that in terms of f we may characterize sets in the independence ideal of φ as follows: J is in this ideal if and only if $(K', K'') \subseteqq J$ implies $f(K') = f(K'')$, where (K', K'') is the symmetric difference $K' \cdot \widetilde{K''} + \widetilde{K'} \cdot K''$. For, suppose φ is independent of J, and $(K', K'') \subseteqq J$. Then $f(K') = \varphi(x'_\alpha \,|\, \alpha \in I)$, where $x'_\alpha = 1$ for $\alpha \in K'$, and $x'_\alpha = 0$ for $\alpha \notin K'$, and $f(K'') = \varphi(x''_\alpha \,|\, \alpha \in I)$, where $x''_\alpha = 1$ for $\alpha \in K''$, and $x''_\alpha = 0$ for $\alpha \notin K''$. Hence $x'_\alpha = x''_\alpha$ for all α not in (K', K''), i.e. for all α not contained in a subset of J, so that $\varphi(x'_\alpha \,|\, \alpha \in I) = \varphi(x''_\alpha \,|\, \alpha \in I)$, or $f(K') = f(K'')$. Conversely, suppose that $(K', K'') \subseteqq J$ implies $f(K') = f(K'')$, and that $(x'_\alpha \,|\, \alpha \in I)$ and $(x''_\alpha \,|\, \alpha \in I)$ are two sequences for which $x'_\alpha = x''_\alpha$ for $\alpha \notin J$. Let K' be the set of α's for which $x'_\alpha = 1$, and K'' the set of α's for which $x''_\alpha = 1$. Then it is easy to see that $(K', K'') \subseteqq J$, so that

$$\varphi(x'_\alpha \,|\, \alpha \in I) = f(K') = f(K'') = \varphi(x''_\alpha \,|\, \alpha \in I);$$

i.e. φ is independent of J.

It is now easy to define a set function f such that the independence ideal of the corresponding φ is precisely $\mathcal{T}$: we define $f(K) = 1$ if $K \in \mathcal{T}$ and $f(K) = 0$ otherwise. To prove that f has the required property, first let J be any set in $\mathcal{T}$, and K' and K'' two sets for which $(K', K'') \subseteqq J$. Then, of course, $(K', K'') \in \mathcal{T}$; since $K' \subseteqq K'' + (K', K'')$ and $K'' \subseteqq K' + (K', K'')$, if either K' or K'' is in $\mathcal{T}$, so is the other one; hence $f(K') = f(K'')$. Next suppose $J \notin \mathcal{T}$, and take $K' = J$, $K'' = \theta$. Then $(K', K'') = J$, but $f(K') = f(J) = 0$ and $f(K'') = f(\theta) = 1$. This completes the proof of 7.3.

We conclude this section with the following obvious observations:

7.4. *The function $\varphi(\xi) = \varphi(x_\alpha \,|\, \alpha \in I)$ is a constant if and only if I is an independence set of φ, i.e. if and only if* **I** *is the independence ideal of φ.*

Returning to the discussion at the beginning of §7: the independence of φ from each one of its variables x_α means merely that every one-element set (α) is an independence set of φ; i.e. (cf. 7.2 and 6.1) every finite set is in the independence ideal of φ. Owing to the above and to 7.3 this implies that φ is a constant if and only if I is finite.

8. Commutativity

Throughout the remainder of this chapter we shall use the following notations. S is a fixed Hausdorff space, I is an arbitrary set of indices; for each $\alpha \in I, S_\alpha = S$; Ξ is the sequence space of elements $\xi = (x_\alpha \,|\, \alpha \in I)$, $x_\alpha \in S_\alpha = S$. As in 5.4 we write $C(\xi) = C(x_\alpha \,|\, \alpha \in I)$ for the closure (in S) of the set of points x_α, $\alpha \in I$.

We define:

8.1.1. A sequence ξ is bounded if and only if $C(\xi)$ is a compact set.

And we prove:

8.1.2. *A sequence ξ is bounded if and only if there exists a compact set C, such that $x_\alpha(\xi) \in C$ for all $\alpha \in I$.*

PROOF. Necessity: Put $C = C(\xi)$.
Sufficiency: C is closed by 5.1, so $C(\xi) \subseteqq C$. Now $C(\xi)$ is a closed subset of a compact set; consequently it is compact by 1.5.

We are now in a position to specify the set $\Xi^* \subseteqq \Xi$ which will be the domain of all our functions $\varphi(\xi) = \varphi(x_\alpha \,|\, \alpha \in I)$:

8.1.3. Ξ^* is the set of all bounded sequences $\xi \in \Xi$.

This Ξ^* fulfills our condition 7.1.1: Under the assumptions of 7.1.1 every $x_\alpha^0 = x_\alpha(\xi_0) \in C(\xi_1) + C(\xi_2)$; hence ξ_0 is bounded if ξ_1, ξ_2 are, owing to 8.1.2.

8.2. We pass now to considering another family of functions, to be denoted by ψ. To this end we introduce a new index set J (which has no relationship with I), and assume that a compact set $E_\beta \subseteqq S$ is given for every $\beta \in J$. We then put $S_\beta = E_\beta$ for all $\beta \in J$.
The space of all sequences $\eta = (y_\beta \,|\, \beta \in J)$ $(y_\beta \in S_\beta = E_\beta \subseteqq S)$ will be called* H (this corresponds to the Ξ of 5.4). We consider the functions $\psi(\eta) = \psi(y_\beta \,|\, \beta \in J)$ which are defined for all $\eta \in \mathrm{H}$ (i.e., in forming the H^* which corresponds to the Ξ^* of 5.4, we put $\mathrm{H}^* = \mathrm{H}$). Thus our condition 7.1.1 is automatically fulfilled.
We assume furthermore that our functions $\psi(\eta) = \psi(y_\beta \,|\, \beta \in J)$ assume values from S, and are continuous in their entire domain ($y \in \mathrm{H}$) with respect to the topology of 5.6 (with J in place of I).
We prove

8.2.1. *Consider a function ψ (and J, E_β) as described above, and a (double) sequence $\zeta = (z_{\alpha\beta} \,|\, \alpha \in I, \beta \in J)$ with $z_{\alpha\beta} \in E_\beta$ for all $\alpha \in I$, $\beta \in J$. Denote by $\psi[\zeta]$ the (simple) sequence $(\psi(z_{\alpha\beta} \,|\, \beta \in J) \,|\, \alpha \in I)$. Then $\psi[\zeta]$ is bounded.*

PROOF. The domain H of $\psi(\eta)$ is compact, since all E_β are: We have the situation of 5.5, and so we may apply Tychonoff's theorem. Now the range of $\psi(\eta)$ is the image of H by the continuous mapping $\eta \to \psi(\eta)$; hence it is also compact by 5.3. Consequently the sequence $\psi[\zeta] = (\psi(z_{\alpha\beta} \,|\, \beta \in J) \,|\, \alpha \in I)$ is bounded by 8.1.2.

8.3. Let ψ (and J, E_β) be as described above, and assume that a (double) sequence $\zeta = (z_{\alpha\beta} \,|\, \alpha \in I, \beta \in J)$ with

8.3.1. $z_{\zeta\beta} \in E_\beta$ for all $\alpha \in I$, $\beta \in J$,

8.3.2. $\varphi(z_{\alpha\beta} \,|\, \alpha \in I) \in E_\beta$ for all $\beta \in J$,

is given. Denote by $\varphi[\zeta]$ the (simple) sequence $(\varphi(z_{\alpha\beta} \,|\, \alpha \in I) \,|\, \beta \in J)$ and by $\psi[\zeta]$ (as in 8.2.1) the (simple) sequence $(\psi(z_{\alpha\beta} \,|\, \beta \in J) \,|\, \alpha \in I)$. Combining 8.3.1, 8.3.2 and 8.2.1 we see that the expression

$$\varphi(\psi[\zeta]) = \varphi(\psi(z_{\alpha\beta} \,|\, \beta \in J) \,|\, \alpha \in I), \tag{8.3.3}$$

as well as the expression

$$\psi(\varphi[\zeta]) = \psi(\varphi(z_{\alpha\beta} \,|\, \alpha \in I) \,|\, \beta \in J), \tag{8.3.4}$$

is meaningful. We now define:

---

*A roman cap H is used for the Greek cap eta.

8.3.5. Under the conditions stated above (for φ, ψ and J, E_β) we say that φ is commutative with ψ if for all (double) sequences $\zeta = (z_{\alpha\beta} \mid \alpha \in I,\ \beta \in J)$ fulfilling 8.3.1, 8.3.2, the expression of (8.3.3) is equal to that one of (8.3.4). I.e.

$$\varphi(\psi[\zeta]) = \psi(\varphi[\zeta]).$$

We now proceed to derive two easy consequences of commutativity.

8.4. *If φ is commutative with all (continuous) functions ψ of one variable (i.e. with a one-element set J—for details concerning φ, ψ, and J, E_β cf.* 8.2)*—then for every bounded sequence ξ*

$$\varphi(\xi) \in C(\xi).$$

PROOF. Let $\xi = (x_\alpha \mid \alpha \in I)$ be any bounded sequence, and write $\varphi(\xi) = x^0$. If $x^0 \notin C(\xi)$, let E be the set $C(\xi) + (x^0)$ (then E is a compact set), and let x' be an arbitrary point of S, $x' \neq x^0$. We define

$$\psi(x) = x \quad \text{for } x \in C(\xi), \qquad \psi(x^0) = x'.$$

Then $\psi(x)$ is a continuous function with domain E, so that we should have $\varphi\psi = \psi\varphi$. The following computation shows that this is not the case and therefore contradicts the hypothesis that $x^0 \notin C(\xi)$.

$$\varphi\psi = \varphi(\psi(x^0_\alpha) \mid \alpha \in I) = \varphi(x^0_\alpha \mid \alpha \in I) = x^0;$$
$$\psi\varphi = \psi(\varphi(x^0_\alpha \mid \alpha \in I)) = \psi(x^0) = x'.$$

We observe that owing to this result, for the φ in question 8.3.1 implies 8.3.2.

8.5. *If φ is commutative with all (continuous) functions of two variables (i.e. with a two-element J; for details concerning φ, ψ, and J cf.* 8.2)*, then the independence ideal of φ is a prime ideal.*

PROOF. Let $\mathcal{T}$ be the independence ideal of φ. We will establish its prime ideal character with the aid of 6.2.1. Now if φ were a constant, it could not fulfill the above commutativity condition (to see this, choose ψ as another constant $\neq \varphi$); hence $\mathcal{T} \neq \mathbf{I}$ (cf. 7.4). Consequently we must only show that for any set $I' \subseteqq I$, $\mathcal{T}$ contains at least one of the sets I' and $\widetilde{I'}$.

Consider therefore an $I' \subseteqq I$ and let p and q be arbitrary points in S, $p \neq q$. Write $y_\alpha = p$ for $\alpha \in I'$ and $y_\alpha = q$ for $\alpha \notin I'$. It follows from 8.4 that the value of $\varphi(y_\alpha \mid \alpha \in I)$ is either p or q. Suppose that it is p.

Consider a bounded sequence $\xi = (x_\alpha \mid \alpha \in I)$. Define another sequence $\xi^* = (x^*_\alpha \mid \alpha \in I)$ by

$$x^*_\alpha = x_\alpha \text{ for } \alpha \in I', x_\alpha = q \text{ for } \alpha \notin I'. \tag{8.5.1}$$

ξ^* is also bounded, since $C(\xi^*) \subseteqq C(\xi) + (q)$.

Now put $E' = C(\xi)$, $E'' = (p, q)$, and define a two-variable function $\psi(x, y)$, where x has the domain E' and y the domain E'', as follows:

$$\psi(x, y) = x \quad \text{for } y = p, \qquad \psi(x, y) = q \quad \text{for } y = q.$$

This $\psi(x, y)$ is clearly continuous, so φ commutes with it. We have therefore (cf. 8.3, and the remark at the end of 8.4):

$$\varphi(\psi(x_\alpha, y_\alpha) \mid \alpha \in I) = \psi(\varphi(x_\alpha \mid \alpha \in I), \varphi(y_\alpha \mid \alpha \in I)).$$

The left-hand side is obviously equal to $\varphi(x^*_\alpha \mid \alpha \in I)$, while the right-hand side is equal to $\psi(\varphi(x_\alpha \mid \alpha \in I), p)$, i.e. to $\varphi(x_\alpha \mid \alpha \in I)$. So

$$\varphi(x_\alpha \mid \alpha \in I) = \varphi(x^*_\alpha \mid \alpha \in I). \tag{8.5.2}$$

Combining (8.5.1), (8.5.2) we see that $\widetilde{I}'$ is an independence set of φ, i.e. that $\widetilde{I}' \in \mathcal{T}$.

So $\varphi(y_\alpha \mid \alpha \in I) = p$ implies $\widetilde{I}' \in \mathcal{T}$. Similarly $\varphi(y_\alpha \mid \alpha \in I) = q$ implies $I' \in \mathcal{T}$. Hence either I' or $\widetilde{I}'$ belongs to $\mathcal{T}$ and the proof is completed.

9. Limit functions

9.1. A function $\varphi(\xi)$ defined for all bounded sequences ξ and commutative with all continuous functions ψ (see 8.3) we shall call a *limit function* or simply a *limit.* It follows from 8.4 and 8.5 that if $\varphi(\xi)$ is a limit $\varphi(\xi) \in C(\xi)$ for all bounded ξ, and the independence ideal of φ is a prime ideal. In this section we shall be concerned mostly with the existence of limits. To save detailed explanations in each special case, we assume that all the function symbols $(\varphi(x_\alpha \mid \alpha \in I), \psi(x_\beta \mid \beta \in J))$ we shall write are meaningful: in other words that the arguments written lie in the domain of the function considered.

9.2. *The set $\Lambda \subseteqq \Phi$ of all functions $\varphi(\xi)$, which are defined for all bounded sequences ξ and are such that $\varphi(\xi) \in C(\xi)$, is a compact set.*

PROOF. This is merely a restatement of Tychonoff's theorem 5.5 with $\Xi(\xi) = C(\xi)$. (For the notation see 5.6.) We remark that it follows from 9.1 (see also 8.4) that the set of all limit functions is a subset of the compact space Λ.

9.3. *The set of all limit functions is a closed subset of* Λ.

PROOF. We shall give the proof in four steps.

(i) Let ψ_0 (and J, E_β) be as described in 8.2. (In particular, ψ_0 is continuous.) Consider an arbitrary $\varphi \in \Lambda$ and a fixed sequence $\zeta^0 = (z^0_{\alpha\beta} \mid \alpha \in I,\ \beta \in J)$ with the property 8.3.1, i.e.

$$z^0_{\alpha\beta} \in E_\beta \text{ for all } \alpha \in I, \beta \in J. \tag{9.3.1}$$

Then 8.3.2 is also true, owing to the definition of Λ in 9.2. Consider now the two expressions of (8.3.3), (8.3.4), i.e.

$$\varphi(\psi_0(z^0_{\alpha\beta} \mid \beta \in J) \mid \alpha \in I), \tag{9.3.2}$$

$$\psi_0(\varphi(z^0_{\alpha\beta} \mid \alpha \in I) \mid \beta \in J). \tag{9.3.3}$$

We repeat: $\psi_0, \zeta^0 = (z^0_{\alpha\beta} \mid \alpha \in I,\ \beta \in J)$ are fixed; φ is variable. Let us fix our attention on a $\varphi_0 \in \Lambda$ and let φ vary over Λ.

(ii) (9.3.2) is a continuous function of φ (in Λ, at φ_0). Let N be a neighborhood of $\varphi_0(\psi_0(z^0_{\alpha\beta} \mid \beta \in J) \mid \alpha \in I)$ (in S). Then (by 5.4, remembering 5.6) $\varphi(\psi_0(z^0_{\alpha\beta} \mid \alpha \in I) \mid \beta \in J) \in N$ defines a neighborhood of φ_0 (in Φ, hence also in Λ).

(iii) (9.3.3) is a continuous function of φ (in Λ, at φ_0). Let N be a neighborhood of $\psi_0(\varphi_0(z^0_{\alpha\beta} \mid \alpha \in I) \mid \beta \in J)$ (in S). Then, owing to the continuity of ψ_0 (by 5.4, remembering 8.2), we can find a finite number of elements $\beta_1, \dots, \beta_n$ of J, and open sets $N_1, \dots, N_n$ (in S), such that N_i contains $\varphi_0(z^0_{\alpha\beta_i} \mid \alpha \in I)$ for $i = 1, \dots, n$, with the following property: If the sequence $(y_\beta \mid \beta \in J)$ (with $y_\beta \in E_\beta$ for all $\beta \in J$) fulfills $y_{\beta_i} \in N_i$ for $i = 1, \dots, n$, then $\psi_0(y_\beta \mid \beta \in J) \in N$. Now (by 5.4, remembering 5.6) $\varphi(z^0_{\alpha\beta_i} \,|\, \alpha \in I) \in N_i$ for all $i = 1, \dots, n$ defines a neighborhood of φ_0 (in Φ, hence also in Λ). And this condition implies

$$\psi_0(\varphi(z^0_{\alpha\beta} \,|\, \alpha \in I) \,|\, \beta \in J) \in N$$

(put $y_\beta = \varphi(z^0_{\alpha\beta} \,|\, \alpha \in I)$ for all $\beta \in J$).

(iv) We see from (ii), (iii), that for each ψ_0, $\zeta^0 = (z^0_{\alpha\beta} \,|\, \alpha \in I,\ \beta \in J)$ the set of all $\varphi \in \Lambda$ with $\varphi(\psi_0(z^0_{\alpha\beta} \,|\, \beta \in J) \,|\, \alpha \in I) = \psi_0(\varphi(z^0_{\alpha\beta} \,|\, \alpha \in I) \,|\, \beta \in J)$ is defined by the equality of two continuous functions (in the closed set Λ, cf. 9.2 and 5.1), and is consequently closed. The intersection of all these closed sets, for all possible $\psi_0, \zeta^0 = (z^0_{\alpha\beta} \,|\, \alpha \in I, \beta \in J)$, is therefore also closed. But this is precisely the set of all limit functions.

This completes the proof of 9.3.

9.4. *If I' is any subset of I, the set of all functions φ which are independent of I' is a closed set.*

PROOF. If $(x'_\alpha \,|\, \alpha \in I)$ and $(x''_\alpha \,|\, \alpha \in I)$ are any two bounded sequences, then the set of all functions φ for which $\varphi(x'_\alpha \,|\, \alpha \in I) = \varphi(x''_\alpha \,|\, \alpha \in I)$ is the set of points of equality of two continuous functions and is therefore closed. Hence the intersection of all these sets, formed for sequences $(x'_\alpha \,|\, \alpha \in I)$ and $(x''_\alpha \,|\, \alpha \in I)$ for which $x'_\alpha = x''_\alpha$ when $\alpha \notin I'$, is a closed set, and this is equivalent to the statement of 9.4.

9.5. *If $\Delta(I')$ is the set of all limit functions φ which is independent of I', then*

9.5.1. *$\Delta(I')$ is a closed subset of the compact set Λ;*

9.5.2. *$\Delta(I')$ is empty if and only if $I' = I$;*

9.5.3. *$I' \subseteqq I''$ implies $\Delta(I') \supseteqq \Delta(I'')$.*

PROOF. For 9.5.1: Immediate by 9.2, 9.3, 9.4.

For 9.5.2: No limit function has I for an independence set, owing to 9.1 and 8.5, so that $\Delta(I) = \theta$. Assume now that $I' \neq I$. Then we may find $\alpha_0 \notin I'$. We define $\varphi(\xi) = x_{\alpha_0}(\xi)$, and assert that φ is a limit function independent of I'. The independence property is obvious; and if $\psi_0(y_\beta \,|\, \beta \in J)$ is any function (not necessarily continuous) and $z^0_{\alpha\beta} \in S$ for all $\alpha \in I$ and $\beta \in J$, then we have

$$\varphi(\psi_0(z^0_{\alpha\beta} \,|\, \beta \in J) \,|\, \alpha \in I) = \psi_0(z^0_{\alpha_0,\beta} \,|\, \beta \in J)$$

and

$$\psi_0(\varphi(z^0_{\alpha\beta} \,|\, \alpha \in I) \,|\, \beta \in J) = \psi_0(z^0_{\alpha_0,\beta} \,|\, \beta \in J),$$

so that φ commutes with ψ_0 and is therefore a limit function.

For 9.5.3: This is a consequence of the fact that the collection of sets of which a function is independent is an ideal (see 7.2).

We have already seen that the independence ideal of a limit is a prime ideal. The following two theorems establish the existence of rather general kinds of limits and prove the converse of this statement.

9.6. *If $\mathcal{T}$ is any ideal, $\mathcal{T} \neq \mathbf{I}$, then there exists a limit function $\varphi(\xi)$ with independence ideal $\mathfrak{P}$ such that $\mathcal{T} \subseteqq \mathfrak{P}$.*

PROOF. Using the notation and results of 9.5 we observe that to every $I' \subseteqq \mathcal{T}$ there corresponds a nonempty closed set $\Delta(I')$ of the compact space of all limit functions. (See 9.5.1, 9.5.2.) Moreover, if $\Delta(I'), \ldots, \Delta(I^n)$ is any finite number of these sets, then 9.5.3 implies (since $I^1 + \cdots + I^n$ is also in $\mathcal{T}$) that

$$\prod_{j=1}^{n} \Delta(I^j) \supseteqq \Delta(I^1 + \cdots + I^n)$$

is not empty. It follows from 5.2 that there is an element $\varphi(\xi)$ common to all $\Delta(I')$ $(I' \in \mathcal{T})$, and this is what was to be proved.

9.6.1. We observe that 9.6 has the consequence that every ideal $\mathcal{T} \neq \mathbf{I}$ is contained in a prime ideal.

(We obtained this result as a byproduct of 9.6 and 8.5. Various simple direct proofs, however, are known. Cf. loc. cit. 6.3.)

9.7. *Every prime ideal is the independence ideal of at least one limit function.*

PROOF. According to 9.6 we may find a limit function whose independence ideal $\mathcal{T}$ contains the given prime ideal $\mathfrak{P}$. Now $\mathcal{T} \neq \mathbf{I}$ (cf. 8.5) and $\mathfrak{P}$ possesses the property 6.2.2; consequently $\mathcal{T} = \mathfrak{P}$.

9.8. The notion of *limit* as we introduced it in this section is a modification and a very wide generalization of the *general limit* of Banach and Mazur. (The latter will be discussed and correlated to our limit in §12.)

Our existence proof in 9.2–9.6 is based on an idea of Kakutani and Ulam, which these authors used in a simplified derivation of the above-mentioned results of Banach and Mazur, and of certain generalizations of the same.

The essentially new aspect in our procedure is in the introduction of the general notion of independence, the use of set theoretical ideals and prime ideals, and the simultaneous use of all continuous functions ψ.

10. Uniqueness

10.1. A prime ideal $\mathfrak{P}$ is of the *first kind* if there is an $\alpha \in I$ such that no set of $\mathfrak{P}$ contains α; otherwise we say that $\mathfrak{P}$ is of the *second kind.*

10.2. We define:

10.2.1. Given an $\alpha \in I$, let $\mathfrak{P}_\alpha$ be the system of all $I' \subseteqq I$ with $\alpha \notin I'$.

10.2.2. Let $\mathfrak{F}$ be the system of all finite $I' \subseteqq I$.

Clearly, $\mathfrak{P}_\alpha$, $\mathfrak{F}$ are ideals, $\mathfrak{P}_\alpha$ is even a prime ideal (e.g. by 6.2.3), and $\mathfrak{F} \neq \mathbf{I}$ if and only if I is infinite. One verifies easily that $\mathfrak{F}$ is not a prime ideal.

We now prove:

10.2.3. *A prime ideal $\mathfrak{P}$ is of the first kind, if and only if it is a $\mathfrak{P}_\alpha$, $\alpha \in I$.*

10.2.4. *A prime ideal $\mathfrak{P}$ is of the second kind, if and only if it is $\supseteqq \mathfrak{F}$.*

10.2.5. *Prime ideals of the second kind exist if and only if I is infinite.*

PROOF. For 10.2.3: Sufficiency: We know that $\mathfrak{P}_\alpha$ is a prime ideal.

Necessity: If $\mathfrak{P}$ is of the first kind, then an $\alpha \in I$ exists such that $I' \in \mathfrak{P}$ implies $\alpha \notin I'$. I.e. $\mathfrak{P} \subseteqq \mathfrak{P}_\alpha$. Hence by 6.2.2 $\mathfrak{P} = \mathfrak{P}_\alpha$.

For 10.2.4: That $\mathfrak{P}$ is of the second kind, means that for every $\alpha \in I$ an $I' \in \mathfrak{P}$ with $\alpha \in I'$, i.e. $(\alpha) \subseteqq I'$, exists—that is, $(\alpha) \in \mathfrak{P}$. Since $\mathfrak{P}$ is an ideal, this means that $\mathfrak{P}$ must contain all finite sets $I' \subseteqq I$, i.e. $\mathfrak{P} \supseteqq \mathfrak{F}$.

For 10.2.5: Owing to 10.2.4 and 9.6 prime ideals of the second kind exist if and only if $\mathfrak{F} \neq \mathbf{I}$, i.e. if and only if I is infinite.

The above results indicate that the prime ideals of the first kind are equivalents of the elements of I, while those of the second kind are something like an extension of I. The idea of extending a system by considering its prime ideals is a fundamental one in modern algebra and algebraic geometry; it was introduced into set theory and topology by Stone. (Cf. loc. cit. 6.3.)

10.3. In this section we shall prove that to every prime ideal there corresponds *exactly one* limit whose independence ideal it is (see 9.7). Our proof will be valid for prime ideals of both kinds, but we shall first observe that this fact, as well as another one connected with it, is trivial for prime ideals of the first kind.

10.3.0. *Let $\mathfrak{P} = \mathfrak{P}_{\alpha_0}$ be a prime ideal of the first kind. Then we have*:

10.3.1. *There exists one and only one limit function with the independence ideal $\mathfrak{P} : \varphi_{\mathfrak{P}}(\xi) = x_{\alpha_0}(\xi)$.*

10.3.2. *In fact, the above $\varphi_{\mathfrak{P}}(\xi)$ is the only function $\varphi(\xi)$ with an independence ideal containing $\mathfrak{P}$ and with $\varphi(\xi) \in C(\xi)$.*

10.3.3. *The function $\omega(\xi)$ has an independence ideal containing $\mathfrak{P}$ if and only if $\omega(\xi) = f(\varphi_{\mathfrak{P}}(\xi))$.*

PROOF. For 10.3.3: That the independence ideal of $\omega(\xi)$ contains $\mathfrak{P} = \mathfrak{P}_{\alpha_0}$ is obviously equivalent to the statement that the set $I - (\alpha_0)$ belongs to it, i.e. that $I - (\alpha_0)$ is an independence set of $\omega(\xi)$. This means $\omega(\xi) = f(x_{\alpha_0})$ when $\xi = (x_\alpha \,|\, \alpha \in I)$, i.e. $\omega(\xi) = f(x_{\alpha_0}(\xi))$, or, with our $\varphi_{\mathfrak{P}}(\xi)$ of 10.3.1, $\omega(\xi) = \mathfrak{f}(\varphi_{\mathfrak{P}}(\xi))$.

For 10.3.2: If $\xi = (x_\alpha \,|\, \alpha \in I)$, put $\xi^*_{\alpha_0} = (x_{\alpha_0} \,|\, \alpha \in I)$. Now $\varphi_{\mathfrak{p}}(\xi) = \varphi_{\mathfrak{p}}(\xi^*_0) = x_{\alpha_0}$. Hence by 10.3.3 $\varphi(\xi) = \varphi(\xi^*_{\alpha_0})$; hence with our present assumptions

$$\varphi(\xi) \in C(\xi^*_{\alpha_0}) = (x_{\alpha_0}), \qquad \varphi(\xi) = x_{\alpha_0} = x_{\alpha_0}(\xi) = \varphi_{\mathfrak{P}}(\xi).$$

$$\varphi(\xi) = \varphi(\xi^*_{\alpha_0}) \in C(\xi^*_{\alpha_0}) = (x_{\alpha_0}).$$

For 10.3.3: Immediate by 10.3.2 and 8.4.

10.4. We now proceed to analyze the situation for arbitrary prime ideals $\mathfrak{P}$. The discussions which follow provide also the basis for the considerations of the next section. Together with those they could be used for an alternative development of the theory of limits.

We begin by considering ideals $\mathcal{T}$ which are not necessarily prime.

If $\xi = (x_\alpha \,|\, \alpha \in I)$ is a bounded sequence, and $I' \subseteqq I$, then we write $C_{I'}(\xi) = C(x_\alpha \,|\, \alpha \in \widetilde{I'})$ for the closure (in S) of the set of points x_α, $\alpha \in \widetilde{I'}$. The following statements are obvious:

10.4.1. *$C_{I'}(\xi)$ is a closed subset of the compact set $C_\theta(\xi) = C(\xi)$.*

10.4.2. *$C_{I'}(\xi)$ is empty if and only if $I' = I$.*

10.4.3. *$I' \subseteqq I''$ implies $C_{I'}(\xi) \supseteqq C_{I''}(\xi)$.*

From these we conclude:

10.4.4. *Let $\mathcal{T}$ be an arbitrary ideal. The set*

$$L^{\mathcal{T}}(\xi) = \prod_{I' \in \mathcal{T}} C_{I'}(\xi)$$

is empty if and only if $\mathcal{T} = \mathbf{I}$.

PROOF. Sufficiency: Assume $\mathcal{T} = \mathbf{I}$. Then $L^{\mathcal{T}}(\xi) \subseteqq C_I(\xi) = \theta$ by 10.4.2.

Necessity: Assume $\mathcal{T} \neq \mathbf{I}$. Then $L^{\mathcal{T}}(\xi) \neq \theta$ follows from 10.4.1–10.4.3 in literally the same way as 9.6 from 9.5.1–9.5.3.

10.5. Consider now a prime ideal $\mathfrak{P}$.

10.5.1. *$L^{\mathfrak{P}}(\xi)$ is a one-element set.*

PROOF. $L^{\mathfrak{P}}(\xi)$ is not empty by 10.4.4. Consequently we must only prove that it cannot have two different elements.

Assume therefore $p, q \in L^{\mathfrak{P}}(\xi)$, $p \neq q$. Choose two disjoint neighborhoods P of p and Q of q. Then $F = L^{\mathfrak{P}}(\xi) \cdot \widetilde{P}$, $G = L^{\mathfrak{P}}(\xi) \cdot \widetilde{Q}$ are closed sets, and

(10.5.1.1) $$p \notin F, q \notin G,$$

(10.5.1.2) $$F + G = L^{\mathfrak{P}}(\xi).$$

Let I' be the set of all $\alpha \in I$ with $x_\alpha \in F$, and I'' the set of all $\alpha \in I$ with $x_\alpha \in G$. Owing to (10.5.1.2) $I' + I'' = I$, so

(10.5.1.3) $$\widetilde{I'} \subseteqq I'', \widetilde{I''} \subseteqq I'.$$

Also, since F, G are closed, and considering (10.5.1.3), we have

(10.5.1.4) $$C_{I'}(\xi) \subseteqq G, C_{I''}(\xi) \subseteqq F.$$

Since $\mathfrak{P}$ is a prime ideal, so $I' \in \mathfrak{P}$ or $\widetilde{I'} \in \mathfrak{P}$, by 6.2.1; hence $I' \in \mathfrak{P}$ or $I'' \in \mathfrak{P}$. We may assume that $I' \in \mathfrak{P}$. Then $L^{\mathfrak{P}}(\xi) \subseteqq C_{I'}(\xi)$; hence (10.5.1.2) and (10.5.1.4) give $G = L^{\mathfrak{P}}(\xi)$. But this contradicts (10.5.1.1).

The proof is thus completed.

10.5.2. *Let $\varphi(\xi)$ be a function with an independence ideal containing $\mathfrak{P}$ and with $\varphi(\xi) \in C(\xi)$. Then $\varphi(\xi) \in L^{\mathfrak{P}}(\xi)$.*

PROOF. Write $\xi = (x_\alpha \mid \alpha \in I)$, and consider an $I' \in \mathfrak{P}$. Choose an $\alpha_0 \in \widetilde{I'}$, and form $\xi' = (x'_\alpha \mid \alpha \in I)$, where

$$x'_\alpha = x_\alpha \quad \text{for } \alpha \in \widetilde{I}, \qquad x'_\alpha = x_{\alpha_0} \quad \text{for } \alpha \in I.$$

Then clearly $C(\xi') = C_{I'}(\xi)$, and $\varphi(\xi) = \varphi(\xi') \in C(\xi') = C_{I'}(\xi)$, i.e. $\varphi(\xi) \in C_{I'}(\xi)$. This is true for all $I' \in \mathfrak{P}$ so $\varphi(\xi) \in L^{\mathfrak{P}}(\xi)$.

We are now in a position to prove:

10.5.3. *There exists one and only one limit function with the independence ideal $\mathfrak{P}$: $\varphi_{\mathfrak{P}}(\xi)$.*

10.5.4. *In fact, the above $\varphi_{\mathfrak{P}}(\xi)$ is the only function $\varphi(\xi)$ with an independence ideal containing $\mathfrak{P}$ and with $\varphi(\xi) \in C(\xi)$.*

10.5.5. *$L^{\mathfrak{P}}(\xi)$ has the unique element $\varphi_{\mathfrak{P}}(\xi)$.*

PROOF. For 10.5.4, 10.5.5: These are just restatements of the combination of 10.5.1, 10.5.2.

For 10.5.3: Immediate by 10.5.4 and 8.4.

The analogues of 10.3.1 and 10.3.2 are thereby established. That one of 10.3.3 will be considered in the next section (cf. 11.6).

10.6. The definition of a limit function given in 9.1 can be weakened as follows.

10.6.1. *In order that a function $\varphi(\xi)$ defined for all bounded sequences ξ be a limit function, it is sufficient (and of course also necessary, cf.* 9.1*) that it be commutative with all (continuous) functions ψ of two variables (cf.* 8.5*).*

PROOF. Under these conditions the independence ideal of $\varphi(\xi)$ is a prime ideal by 8.5, say $\mathfrak{P}$, and $\varphi(\xi) \in C(\xi)$ by 8.4. Hence $\varphi(\xi)$ coincides with the limit function $\varphi_{\mathfrak{P}}(\xi)$ by 10.5.3, 10.5.4.

10.7. We conclude this section by showing that restricting our attention to bounded sequences was necessary.

10.7.1. *Let S be a topological space, and C a closed subset of S. Assume that for every index set I and for every ideal $\mathcal{T} \neq \mathbf{I}$ of subsets of I there exists a function $\varphi(\xi) = \varphi(x_\alpha \mid \alpha \in I)$ defined for all sequences $\xi = (x_\alpha \mid \alpha \in I)$ with $C(\xi) \subseteqq C$ (i.e. $x_\alpha \in C$ for all $\alpha \in I$), and with the following properties:*

(10.7.1.1) *The independence ideal of φ contains $\mathcal{T}$.*

(10.7.1.2) $$\varphi(\xi) \in C(\xi).$$

Then C is compact.

PROOF. Let $\mathfrak{B}$ be an arbitrary family of closed sets $F \subseteqq S$, such that for $F_1, \ldots, F_n \in \mathfrak{B}$ always $F_1 \cdot \cdots \cdot F_n \neq \theta$. By adding all these sets $F_1 \cdot \cdots \cdot F_n$ to $\mathfrak{B}$ we see that there is no loss of generality in assuming that $F, G \in \mathfrak{B}$ imply $F \cdot G \in \mathfrak{B}$. And still $\theta \notin \mathfrak{B}$.

For each $F \in \mathfrak{B}$ select an element x_F^0 of F.

Consider a family $\mathfrak{B}' \subseteqq \mathfrak{B}$ with this property:

10.7.1.3. *There exists an $F_0 \in \mathfrak{B}$ such that $F \in \mathfrak{B}'$ implies $x_F^0 \notin F_0$.*

The set $\mathcal{T}$ of all such $\mathfrak{B}'$ is an ideal of subsets of $\mathfrak{B}$: That $\mathfrak{B}' \in \mathcal{T}$ and $\mathfrak{B}'' \subseteqq \mathfrak{B}'$ imply $\mathfrak{B}'' \in \mathcal{T}$ is clear. And if $\mathfrak{B}', \mathfrak{B}'' \in \mathcal{T}$, then $\mathfrak{B}' + \mathfrak{B}'' \in \mathcal{T}$ because if 10.7.1.3 holds for $\mathfrak{B}'$ with F_0', and for $\mathfrak{B}''$ with F_0'', then it holds for $\mathfrak{B}' + \mathfrak{B}''$ with $F_0' \cdot F_0''$. Furthermore $x_F^0 \in F$ excludes that 10.7.1.3 be true for $\mathfrak{B}$ with any F_0; hence $\mathfrak{B} \notin \mathcal{T}$, i.e. $\mathcal{T} \neq \mathbf{I}$.

So we may apply our hypothesis to $I = \mathfrak{B}$ and this $\mathcal{T}$ and obtain a function $\varphi(\xi) = \varphi(x_F \mid F \in \mathfrak{B})$ which fulfills (10.7.1.1), (10.7.1.2). Since $x_F^0 \in F \subseteqq C$, we can form $\varphi(x_F^0 \mid F \in \mathfrak{B})$.

Consider an $F_0 \in \mathfrak{B}$. Let $\mathfrak{B}'$ be the set of all $F \in \mathfrak{B}$ with $x_F^0 \notin F_0$. Then $\mathfrak{B}' \in \mathcal{T}$. Choose a $G_0 \notin \mathfrak{B}'$ and form $\xi' = (x_F' \mid F \in \mathfrak{B})$ with

$$x_F' = x_F^0 \quad \text{for } F \notin \mathfrak{B}', \qquad x_F' = x_{G_0}^0 \quad \text{for } F \in \mathfrak{B}'.$$

Then every x'_F is an x^0_G with $G \notin \mathfrak{B}'$, i.e. with $x^0_G \in F_0$, so every $x'_F \in F_0$. Consequently $C(\xi') \subseteqq F_0$. Now

$$\varphi(x^0_F \mid F \in \mathfrak{B}) = \varphi(x'_F \mid F \in \mathfrak{B}) \in C(\xi') \subseteqq F_0,$$

i.e. $\varphi(x^0_F \mid F \in \mathfrak{B}) \in F_0$. This holds for every $F_0 \in \mathfrak{B}$ so $\varphi(x^0_F \mid F \in \mathfrak{B}) \in \prod_{F_0 \in \mathfrak{B}} F_0$.

So $\prod_{F_0 \in \mathfrak{B}} F_0$ is not empty, and therefore C is compact by 5.2.

11. Convergence

So far the analogies between our limits φ and the classical notion of limit (in topology and in analysis) have been rather formal. We shall show in this section that the connections between these notions are really quite intrinsic and essential.

11.1. We restate the one-to-one correspondence which exists for every index set I between the prime ideals $\mathfrak{P}$ of subsets of I and the limit functions $\varphi(\xi) = \varphi(x_\alpha \mid \alpha \in I)$: $\mathfrak{P}$ being the independence ideal of $\varphi(\xi)$ (cf. 8.4), and $\varphi(\xi)$ the unique $\varphi_{\mathfrak{P}}(\xi)$ (cf. 10.5.3).

11.2. *Let $\mathcal{T}$ be an ideal of subsets of I. Then $L^{\mathcal{T}}(\xi)$ (cf. 10.4.4) is the set of all $\varphi_{\mathfrak{P}}(\xi)$, for all prime ideals $\mathfrak{P} \supseteqq \mathcal{T}$.*

PROOF. Owing to 10.5.5 the above statement means that

$$L^{\mathcal{T}}(\xi) = \sum_{\mathfrak{P} \supseteqq \mathcal{T}} L^{\mathfrak{P}}(\xi).$$

($\mathfrak{P}$ denotes always a prime ideal.) Now $\mathfrak{P} \supseteqq \mathcal{T}$ implies $L^{\mathfrak{P}}(\xi) \subseteqq L^{\mathcal{T}}(\xi)$, and consequently $L^{\mathcal{T}}(\xi) \supseteqq \sum_{\mathfrak{P} \supseteqq \mathcal{T}} L^{\mathfrak{P}}(\xi)$. Hence we must only prove $L^{\mathcal{T}}(\xi) \subseteqq \sum_{\mathfrak{P} \supseteqq \mathcal{T}} L^{\mathfrak{P}}(\xi)$, i.e. this:

For any $p \in L^{\mathcal{T}}(\xi)$ there exists a prime ideal $\mathfrak{P} \supseteqq \mathcal{T}$ with $p \in L^{\mathfrak{P}}(\xi)$.

Observe first that for any $I', I'' \subseteqq I$

$$C_{I' \cdot I''}(\xi) = C_{I'}(\xi) + C_{I''}(\xi) \text{(cf. 10.4)}. \tag{11.2.1}$$

Consider now any $p \in L^{\mathcal{T}}(\xi)$. Let $\mathcal{T}^0$ be the set of all $I' \subseteqq I$ with $p \notin C_{\widetilde{I'}}(\xi)$. $\mathcal{T}^0$ is an ideal by 10.4.3, (11.2.1) (both applied to $\widetilde{I'}, \widetilde{I''}$). Consider the set $\mathcal{T}^*$ of all $I' + I''$, $I' \in \mathcal{T}$, $I'' \in \mathcal{T}^0$. Obviously $\mathcal{T}^*$ is also an ideal, and we have

$$\mathcal{T}^* \supseteqq \mathcal{T}, \mathcal{T}^* \supseteqq \mathcal{T}^0. \tag{11.2.2}$$

$\mathcal{T}^0 = \mathbf{I}$ would imply $I \in \mathcal{T}^0$, i.e. $I = I' + I''$, $I' \in \mathcal{T}$, $I'' \in \mathcal{T}^0$. Then $I'' \supseteqq \widetilde{I'}$; hence $\widetilde{I'} \in \mathcal{T}^0$. Now $I' \in \mathcal{T}$ implies $p \in C_{I'}(\xi)$, while $\widetilde{I'} \in \mathcal{T}^0$ implies $p \notin C_{I'}(\xi)$, which is a contradiction. Thus $\mathcal{T}^* \neq \mathbf{I}$.

Consequently there exists a prime ideal $\mathfrak{P} \supseteqq \mathcal{T}^*$, by 9.6.1. (11.2.2) gives

$$\mathfrak{P} \supseteqq \mathcal{T},\ \mathfrak{P} \supseteqq \mathcal{T}^0. \tag{11.2.3}$$

If $p \notin C_{I'}(\xi)$, then $\widetilde{I'} \in \mathcal{T}^*$ so by (11.2.3) $\widetilde{I'} \in \mathfrak{P}$. This necessitates $I' \notin \mathfrak{P}$, e.g. by 6.2.3. So $I' \in \mathfrak{P}$ implies $p \in C_{I'}(\xi)$. Therefore $p \in L^{\mathfrak{P}}(\xi)$, which, together with (11.2.3), completes the proof.

11.3. We now define: A property $\mathcal{P}(\alpha)$ is true for *almost all* $\alpha \in I$, *modulo* $\mathcal{T}$, if the set of all $\alpha \in I$ for which $\mathcal{P}(\alpha)$ is not true belongs to $\mathcal{T}$.

And further: A bounded sequence $\xi = (x_\alpha \mid \alpha \in I)$ *converges to* p *modulo* $\mathcal{T}$, if for every neighborhood N of p there is $x_\alpha \in N$ for almost all $\alpha \in I$, modulo $\mathcal{T}$.

We observe:

11.3.1. *ξ converges to every p $(\in S)$ modulo* $\mathbf{I}$.

11.3.2. *ξ converges to a unique p $(\in S)$ or to none at all modulo $\mathcal{T}$, if $\mathcal{T} \neq \mathbf{I}$.*

11.3.3. *If $\mathcal{T}' \subseteqq \mathcal{T}''$, then convergence modulo $\mathcal{T}'$ implies convergence modulo $\mathcal{T}''$.*

PROOF. 11.3.1 and 11.3.3 are obvious. For 11.3.2: Assume that $\xi = (x_\alpha \mid \alpha \in I)$ converges to both p and q modulo $\mathcal{T}$, and $p \neq q$. Choose two disjoint neighborhoods P of p and Q of q. Let I' be the set of all $\alpha \in I$ with $x_\alpha \in P$ and I'' the set of those with $x_\alpha \in Q$. Then $I' \cdot I'' = \theta$, and consequently $\widetilde{I'} + \widetilde{I''} = I$.

Now $\widetilde{I'}, \widetilde{I''}$ both belong to $\mathcal{T}$ owing to our hypotheses. Hence $\widetilde{I'} + \widetilde{I''} = I + \mathcal{T}$, $\mathcal{T} \in \mathbf{I}$.

11.4. The above elementary considerations indicate a close analogy between our convergence modulo $\mathcal{T}$ and the classical notion of convergence. Indeed, the two notions clearly coincide, if we choose $\mathcal{T} = \mathfrak{F}$ (cf. 10.2.2).

After these preliminaries we proceed to an exhaustive characterization of convergence modulo $\mathcal{T}$.

11.4.1. *ξ converges to p modulo $\mathcal{T}$, whenever the set $L^{\mathcal{T}}(\xi)$ has the unique element p.*

PROOF. Let N be a neighborhood of p. Then $N \supseteqq L^{\mathcal{T}}(\xi)$, $L^{\mathcal{T}}(\xi) \cdot \widetilde{N} = \theta$; i.e. $\prod_{I' \in \mathcal{T}} C_{I'}(\xi) \cdot \widetilde{N} = \theta$. Since $\widetilde{N}$ is closed, 10.4.1–10.4.3 imply by literally the same argument which was used in proving 9.6 and 10.4.4 that $C_{I'}(\xi) \cdot \widetilde{N} = \theta$ for some $I' \in \mathcal{T}$. I.e. $C_{I'}(\xi) \subseteqq N$ for that $I' \in \mathcal{T}$. Then $\alpha \in \widetilde{I'}$ implies $x_\alpha \in N$; i.e. $x_\alpha \notin N$ implies $\alpha \in I'$. Thus $x_\alpha \in N$ for almost all $\alpha \in I$ modulo $\mathcal{T}$.

This completes the proof.

11.4.2. *ξ converges to $\varphi_{\mathfrak{P}}(\xi)$ modulo $\mathfrak{P}$. ($\mathfrak{P}$ a prime ideal, cf.* 11.1.)

PROOF. Immediate by 11.4.1 and 10.5.5.

11.4.3. *ξ converges to p modulo $\mathcal{T}$ if and only if $L^{\mathcal{T}}(\xi)$ has the unique element p.*

11.4.4. *ξ converges to p modulo $\mathcal{T}$ if and only if $\varphi_{\mathfrak{P}}(\xi) = p$ for all prime ideals $\mathfrak{P} \supseteqq \mathcal{T}$.*

PROOF. 10.4.3 and 10.4.4 are equivalent to each other by 11.2.

Sufficiency of these criteria: In the form 11.4.3 this is merely a restatement of 11.4.1.

Necessity of these criteria: Use the form 11.4.4. Assume that ξ converges to p modulo $\mathcal{T}$. Consider a prime ideal $\mathfrak{P} \supseteqq \mathcal{T}$.

Then ξ converges to p modulo $\mathfrak{P}$ by 11.3.3, and to $\varphi_{\mathfrak{P}}(\xi)$ modulo $\mathfrak{P}$ by 11.4.2. Hence $\varphi_{\mathfrak{P}}(\xi) = p$ by 11.3.2, since $\mathfrak{P} \neq \mathbf{I}$.

11.5. *Every (bounded sequence) ξ converges (to some p) modulo $\mathcal{T}$ if and only if $\mathcal{T}$ is a prime ideal or $\mathcal{T} = \mathbf{I}$.*

PROOF. Sufficiency: Immediate by 11.4.2 and 11.3.1.

Necessity: Assume that $\mathcal{T}$ is not prime and not $\mathbf{I}$. Then there exists by 6.2.1 an $I' \subseteqq I$ with $I' \notin \mathcal{T}$ and $\widetilde{I'} \notin \mathcal{T}$.

Choose two $p, q \in S$ with $p \neq q$. Form $\xi = (x_\alpha \mid \alpha \in I)$ with

$$x_\alpha = p \text{ for } \alpha \in I', x_\alpha = q \text{ for } \alpha \in \widetilde{I}. \tag{11.5.1}$$

$C(\xi) = (p, q)$, so ξ is bounded.

Assume that ξ converges to r modulo $\mathcal{T}$. If $r \neq p$, then let N be a neighborhood of r with $p \notin N$. Then (11.5.1) shows that $x_\alpha \in N$ does not hold for almost all $\alpha \in I$ modulo $\mathcal{T}$. This is impossible; hence $r = p$. Similarly $r = q$. This is a contradiction.

Consequently the above ξ does not converge modulo $\mathcal{T}$.

11.6. We are now able to prove the analogue of 10.3.3. The extra hypothesis of feeble continuity, which we make, is a very weak one, and it is easy to give examples which show that it cannot be omitted when the prime ideal $\mathfrak{P}$ is of the second kind.

The feebly continuous function $\omega(\xi)$ (cf. 5.9) has an independence ideal containing $\mathfrak{P}$ if and only if $\omega(\xi) = f(\varphi_{\mathfrak{P}}(\xi))$.

PROOF. Sufficiency: Obvious.

Necessity: Consider a bounded sequence $\xi = (x_\alpha \mid \alpha \in I)$. Form the sequence $\xi_0 = (\varphi_{\mathfrak{P}}(\xi) \mid \alpha \in I)$. Since $C(\xi_0) = (\varphi_{\mathfrak{P}}(\xi))$, ξ_0 is also bounded. Consider a neighborhood O of $\omega(\xi_0)$. By 5.9 there exists a neighborhood O^0 of $\varphi_{\mathfrak{P}}(\xi)$, such that for every bounded sequence $\xi' = (x'_\alpha \mid \alpha \in I)$ for which $x'_\alpha \in O^0$ for all $\alpha \in I$, we have $\omega(\xi') \in O$.

Now let I' be the set of all $\alpha \in I$ with $x_\alpha \notin O^0$. By 11.3, 11.4.2 $I' \in \mathfrak{P}$. Choose an $\alpha_0 \in \widetilde{I'}$, and form $\xi^* = (x^*_\alpha \mid \alpha \in I)$ where

$$x^*_\alpha = x_\alpha \quad \text{for } \alpha \in \widetilde{I'}, \qquad x^*_\alpha = x_{\alpha_0} \quad \text{for } \alpha \in I'.$$

Since $C(\xi^*) \subseteqq C(\xi)$, ξ^* is bounded. And since the independence ideal of ω contains $\mathfrak{P}$, $\omega(\xi) = \omega(\xi^*)$. Every x^*_α is an x_α with $\alpha \in \widetilde{I'}$, hence with $x_\alpha \in O^0$; i.e. $x^*_\alpha \in O^0$ for all $\alpha \in I$. Therefore $\omega(\xi^*) \in O$; i.e. $\omega(\xi) \in O$.

Since this is true for all neighborhoods O of $\omega(\xi_0)$,

$$\omega(\xi) = \omega(\xi_0). \tag{11.6.1}$$

Now form for any $p \in S$ the sequence $\xi[p] = (p \mid \alpha \in I)$. Since $C(\xi[p]) = (p)$, $\xi[p]$ is bounded. Put

$$f(p) = \omega(\xi[p]). \tag{11.6.2}$$

Clearly $\xi_0 = \xi[\varphi_{\mathfrak{P}}(\xi)]$; hence (11.6.1), (11.6.2) give $\omega(\xi) = f(\varphi_{\mathfrak{P}}(\xi))$, as desired.

11.7. The results of 11.3, 11.4.3, 11.4.4 and 11.5 make the relation of our convergence modulo $\mathcal{T}$ to the classical notion of convergence sufficiently clear. In particular 11.3.1, 11.3.2 and 11.5 show why convergence modulo prime ideals $\mathfrak{P}$ is the fundamental notion on which all others are based through 11.4.2, 11.4.4.

The two following further remarks may also be useful:

11.7.1. ξ converges in the classical sense, i.e. modulo $\mathfrak{F}$, if and only if the $\varphi_{\mathfrak{P}}(\xi)$ have the same value for all prime ideals $\mathfrak{P}$ of the second kind. (Cf. 11.4.4 and 10.2.4.)

11.7.2. Consider a fixed set $I_0 \subseteqq I$. Let $\mathfrak{F}_{I_0}$ be the system of all sets $I' \subseteqq I$ for which $I_0 \cdot I'$ is finite. $\mathfrak{F}_{I_0}$ is clearly an ideal, and $\mathfrak{F}_{I_0} \supseteqq \mathfrak{F}_I = \mathfrak{F}$. Observe that if $I_1 \subseteqq I_2$ and $I_1 \cdot \widetilde{I_2}$ is infinite, then $\mathfrak{F}_{I_1} \supsetneqq \mathfrak{F}_{I_2}$. Consequently when both I_0 and $\widetilde{I_0}$ are infinite, then $\mathfrak{F}_{I_0}$ is greater than $\mathfrak{F}$ but not a maximum ideal, i.e. a prime ideal (cf. 6.3). Somewhat less precisely: $\mathfrak{F}_{I_0}$ is more nearly a prime ideal than $\mathfrak{F}$, but it is nevertheless not a prime ideal.

Consider now a (bounded) sequence $\xi = (x_\alpha \mid \in \alpha I)$. $\xi_{I_0} = (x_\alpha \mid \alpha \in I_0)$ is the general subsequence of ξ. Convergence of the subsequence ξ_{I_0} in the classical sense amounts obviously to convergence of ξ modulo $\mathfrak{F}_{I_0}$. Thus our passing from $\mathfrak{F}$ to prime ideals $\mathfrak{P}$, whereby every bounded sequence becomes convergent, is the absolutely consequent following up of these two convergence-producing processes in analysis: First, the replacement of a sequence by a suitable subsequence, by which any given (bounded) numerical sequence can be made convergent. Second, the "diagonal process" of Cantor-Hilbert, which does the same thing for an enumerably infinite number of given sequences simultaneously. (In these examples it is best to think of the index set $I = (1, 2, 3, \dots)$.)

12. Numerical limits

The subject of this section is a special discussion of the case when S is the set of all real numbers in its usual topology. In this case we speak of *numerical* limits or limit functions. From 12.6 on we shall also assume that $I = (1, 2, \dots)$, but for the time being I is perfectly general.

Observe that a numerical sequence $(x_\alpha \mid \alpha \in I)$ is bounded in the sense of 5.4 if and only if it is bounded in the ordinary sense: when all $|x_\alpha| \leqq c$ for some fixed c.

12.1. We consider functions $\varphi(\xi) = \varphi(x_\alpha \mid \alpha \in I)$ which commute with the following three functions:

$$1, \tag{12.1.1}$$

$$x^2, \tag{12.1.2}$$

$$x + y. \tag{12.1.3}$$

(These are functions of $0, 1, 2$ variables, respectively, playing the role of the $\psi(\eta) = \psi(y_\beta \mid \beta \in J)$ in the theory of limits in the preceding sections. They are everywhere continuous in S.)

12.2. The three commutativity requirements (12.1.1)–(12.1.3) become, in turn, when stated explicitly:

$$\varphi(1 \mid \alpha \in I) = 1; \tag{12.2.1}$$

$$\varphi(x_\alpha^2 \mid \alpha \in I) = (\varphi(x_\alpha \mid \alpha \in I))^2; \tag{12.2.2}$$

$$\varphi(x_\alpha + y_\alpha \mid \alpha \in I) = \varphi(x_\alpha \mid \alpha \in I) + \varphi(y_\alpha \mid \alpha \in I). \tag{12.2.3}$$

12.3. *The conditions of* 12.1 *imply*:

12.3.1. *If all* $x_\alpha \geqq y_\alpha$, *then* $\varphi(x_\alpha \mid \alpha \in I) \geqq \varphi(y_\alpha \mid \alpha \in I)$.

12.3.2. *For every rational number* c

$$\varphi(cx_\alpha \mid \alpha \in I) = c\varphi(x_\alpha \mid \alpha \in I).$$

12.3.3. $\varphi(x_\alpha y_\alpha \mid \alpha \in I) = \varphi(x_\alpha \mid \alpha \in I)\varphi(y_\alpha \mid \alpha \in I)$.

PROOF. For 12.3.1: Applying (12.2.3) to $y_\alpha, x_\alpha - y_\alpha$ (in place of x_α, y_α) and (12.2.2) to $\sqrt{x_\alpha - y_\alpha}$ (in place of x_α—remember $x_\alpha - y_\alpha \geqq 0$) gives

$$\varphi(x_\alpha \mid \alpha \in I) = \varphi(y_\alpha \mid \alpha \in I) + (\varphi(\sqrt{x_\alpha - y_\alpha} \mid \alpha \in I))^2;$$

hence $\varphi(x_\alpha \mid \alpha \in I) \geqq \varphi(y_\alpha \mid \alpha \in I)$.

For 12.3.2: 12.3.2 holds for $c = 1$; if it holds for c_1, c_2, then it holds for $c_1 - c_2$ also (apply 12.2.3 to $c_2 x_\alpha, (c_1 - c_2)x_\alpha$ in place of x_α, y_α); and if it holds for $c_1, c_2 \neq 0$, then it holds for $\frac{c_1}{c_2}$ also (apply it to c_1, x_α and to $c_2, \frac{c_1}{c_2}x_\alpha$ in place of c, x_α). Consequently it holds for all rational c.

For (12.2.3): Use 12.3.2, (12.2.3), (12.2.2); recall $xy = \frac{1}{4}((x+y)^2 - (x-y)^2)$.

12.4. Now the way is open to our main result.

The conditions of 12.1 *are characteristic for a limit function* $\varphi(\xi) = \varphi(x_\alpha \mid \alpha \in I)$.

PROOF. Necessity: Obvious, cf. 9.1.

Sufficiency: Assume that φ satisfies 12.1. By 10.6.1 it suffices to prove that φ commutes with all functions ψ of two variables, i.e.

12.4.1. $\varphi(\psi(x_\alpha, y_\alpha) \mid \alpha \in I) = \psi(\varphi(x_\alpha \mid \alpha \in I), \varphi(y_\alpha \mid \alpha \in I))$ when C, D are two compact sets, such that all $x_\alpha \in C$, all $y_\alpha \in D$, and $\psi(x,y)$ is everywhere defined and continuous within the set $x \in C$, $y \in D$.

C, D are compact, i.e. bounded and closed. Hence C is a closed, finite interval U, from which a finite or enumerably infinite number of open intervals $U^1, U^2, \ldots$ have been removed. Similarly D is a closed, finite interval V, from which a finite or enumerably infinite number of open intervals $V^1, V^2, \ldots$ have been removed.

$\psi(x,y)$ is everywhere defined and continuous within the set $x \in C$, $y \in D$. By linear interpolation in all intervals $U^1, U^2, \ldots$ (for x, still $y \in D$) we can extend $\psi(x,y)$ and make it everywhere defined and continuous within the set $x \in U$, $y \in D$. And then we can apply linear interpolation in all intervals $V^1, V^2, \ldots$ (for y, still $x \in U$), and thereby extend $\psi(x,y)$ and make it everywhere defined and continuous within the set $x \in U$, $y \in V$. I.e.: We can replace C, D by the finite closed intervals U, V. We restate this:

12.4.2. The sets C, D mentioned after 12.4.1 may be assumed to be finite, closed intervals.

Now let a rational $\varepsilon > 0$ be given. By the well-known approximation theorem of Weierstrass, there exists a polynomial $\psi_\varepsilon(x,y)$, such that

12.4.3. $\psi_\varepsilon(x,y) - \varepsilon \leqq \psi(x,y) \leqq \psi_\varepsilon(x,y) + \varepsilon$ for all $x \in C, y \in D$.[4] We can obviously even choose $\psi_\varepsilon(x,y)$ with rational coefficients. Therefore, if 12.4.1 holds for all $\psi_\varepsilon(x,y)$, then the combination of 12.4.3 with (12.2.1), (12.2.3), 12.3.2 and 12.3.1 will extend 12.4.1 to $\psi(x,y)$ also. Hence we have shown:

12.4.4. It suffices to establish 12.4.1 when $\psi(x,y)$ is a polynomial with rational coefficients. But this follows by combining (12.2.1), (12.2.3), 12.3.2, 12.3.3.

Thus the proof is completed.

---

[4]See, for example, R. Courant and D. Hilbert, *Methoden der mathematischen Physik*, Berlin, 1931; vol. 1, pp. 55–57. The approximation theorem is valid for the original compact sets; rather than enter into a discussion of this extraneous fact here we replace it by the extension process above, which is a particularly simple instance of a general theorem.

12.5. The above result simplifies the notion of a numerical limit greatly. Instead of the very general and abstract characterization of 9.1, the three simple conditions of 12.2, i.e. of 12.1, suffice.

The independence ideal of such a limit function $\varphi = \varphi_{\mathfrak{P}}$, the prime ideal $\mathfrak{P}$, is of the first or of the second kind (cf. 10.1). In the first case $\mathfrak{P} = \mathfrak{P}_{\alpha_0}$ for an $\alpha_0 \in I$, and $\varphi(\xi) = \varphi_{\mathfrak{P}}(\xi) = x_{\alpha_0}(\xi)$, i.e. $\varphi(x_\alpha \mid \alpha \in I) = x_{\alpha_0}$ (cf. 10.2.3, 10.3.1). This situation is trivial. The second case, which is the interesting one, is characterized by $(\alpha_0) \in \mathfrak{P}$ for every $\alpha_0 \in I$. I.e. by this condition:

12.5.1. $\varphi(x_\alpha \mid \alpha \in I)$ is independent of x_{α_0} for every $\alpha_0 \in I$.

12.6. Assume now $I = (1, 2, \dots)$. In this case a general notion of limit was introduced by Banach and Mazur.[5] Their postulates differ from our system, which we saw to be equivalent to (12.2.1)–(12.2.3), 12.5.1, and a comparison seems to be of some interest.

12.7. We formulate first a system of postulates which is weaker than our above-mentioned system (12.2.1)–(12.2.3), 12.5.1. It refers to a function $\varphi(x_\alpha \mid \alpha = 1, 2, \dots)$ (we have now $I = (1, 2, \dots)$!) which is defined for all bounded numerical sequences $(x_\alpha \mid \alpha = 1, 2, \dots)$, and it consists of these requirements:

12.7.1. $\varphi(1 \mid \alpha = 1, 2, \dots) = 1$.

12.7.2. If all $x_\alpha \geqq 0$, then $\varphi(x_\alpha \mid \alpha = 1, 2, \dots) \geqq 0$.

12.7.3. $\varphi(x_\alpha + y_\alpha \mid \alpha = 1, 2, \dots) = \varphi(x_\alpha \mid \alpha = 1, 2, \dots) + \varphi(y_\alpha \mid \alpha = 1, 2, \dots)$.

12.7.4. $\varphi(x_\alpha \mid \alpha = 1, 2, \dots)$ is independent of x_{α_0} for every $\alpha_0 = 1, 2, \dots$.

(Indeed: 12.7.1, 12.7.3, 12.7.4 coincide with (12.2.1), (12.2.3), 12.5.1 respectively, and 12.7.2 is a consequence of (12.2.2). To prove the latter, put in (12.2.2) $\sqrt{x_\alpha}$ in place of x_α.)

We derive some consequences of 12.7.1–12.7.3 alone:

12.8. 12.7.1–12.7.3 *imply*:

12.8.1. *If all* $x_\alpha \geqq y_\alpha$, *then*

$$\varphi(x_\alpha \mid \alpha = 1, 2, \dots) \geqq \varphi(y_\alpha \mid \alpha = 1, 2, \dots).$$

12.8.2. *For every real number* c

$$\varphi(cx_\alpha \mid \alpha = 1, 2, \dots) = c\varphi(x_\alpha \mid \alpha = 1, 2, \dots).$$

12.8.3. 12.7.4 *is equivalent to this condition*:

12.8.3.1. *If* $\lim_{\alpha\to\infty} x_\alpha$ *(in the classical sense) exists, then* $\varphi(x_\alpha \mid \alpha = 1, 2, \dots)$ *is equal to it.*

PROOF. For 12.8.1: Apply 12.7.3 to y_α, $x_\alpha - y_\alpha$ (in place of x_α, y_α) and 12.7.2 to $x_\alpha = y_\alpha$ (in place of x_α).

For 12.8.2: Observe that $|x_\alpha| \geqq 0$, $x_\alpha + |x_\alpha| \geqq 0$, and that 12.8.2 holds for x_α if it holds for $|x_\alpha|$ and for $x_\alpha + |x_\alpha|$. I.e.: It suffices to prove 12.8.2 when all $x_\alpha \geqq 0$.

When c is rational, then 12.8.2 is proved literally like 12.3.2. Since all $x_\alpha \geqq 0$, so $\varphi(cx_\alpha \mid \alpha = 1, 2, \dots)$ is a monotone function of c by 12.8.1. Now this is equal to $c\varphi(x_\alpha \mid \alpha = 1, 2, \dots)$ by what we saw above, when c is rational. Hence it is always equal to $c\varphi(x_\alpha \mid \alpha = 1, 2, \dots)$. This completes the proof.

---

[5]Cf., e.g.: S. Banach, *Théorie des opérations linéaires*, Monografje Matematyczne, vol. I, Warsaw (1932) – specifically p. 34.

For 12.8.3: Sufficiency of 12.8.3.1: Assume 12.8.3.1. Assume $x_\alpha = y_\alpha$ for $\alpha \neq \alpha_0$. Then $\lim_{\alpha\to\infty}(x_\alpha - y_\alpha) = 0$; so $\varphi(x_\alpha - y_\alpha \,|\, \alpha = 1,2,\dots) = 0$; i.e. $\varphi(x_\alpha \,|\, \alpha = 1,2,\dots) = \varphi(y_\alpha \,|\, \alpha = 1,2,\dots)$ by 12.7.3.

Necessity of 12.8.3.1: Assume 12.7.4. Consider a sequence $(x_\alpha \,|\, \alpha = 1,2,\dots)$ for which $\lim_{\alpha\to\infty} x_\alpha$ exists, say $\lim_{\alpha\to\infty} x_\alpha = x^0$.

Consider an $\varepsilon > 0$. There exists an α_ε such that $\alpha > \alpha_\varepsilon$ implies $x^0 - \varepsilon \leqq x_\alpha \leqq x^0 + \varepsilon$. Put

$$x_\alpha^\varepsilon = x_\alpha \quad \text{for } \alpha > \alpha_\varepsilon, \qquad x_\alpha^\varepsilon = x^0 \quad \text{for } \alpha \leqq \alpha_\varepsilon.$$

Then always $x^0 - \varepsilon \leqq x_\alpha^\varepsilon \leqq x^0 + \varepsilon$ so by 12.7.1, 12.8.2 and 12.8.1 above, $x^0 - \varepsilon \leqq \varphi(x_\alpha^\varepsilon \,|\, \alpha = 1,2,\dots) \leqq x^0 + \varepsilon$. But repeated application of 12.7.4 (for $\alpha_0 = 1,\dots,\alpha_\varepsilon$) gives $\varphi(x_\alpha \,|\, \alpha = 1,2,\dots) = \varphi(x_\alpha^\varepsilon \,|\, \alpha = 1,2,\dots)$. Hence $x^0 - \varepsilon \leqq \varphi(x_\alpha \,|\, \alpha = 1,2,\dots) \leqq x^0 + \varepsilon$. Since this is true for every $\varepsilon > 0$, therefore $\varphi(x_\alpha \,|\, \alpha = 1,2,\dots) = x^0$. This completes the proof.

12.9. Banach's system obtains from the above 12.7.1–12.7.4, by replacing 12.7.4 by the stronger requirement

$$\varphi(x_{\alpha+1} \,|\, \alpha = 1,2,\dots) = \varphi(x_\alpha \,|\, \alpha = 1,2,\dots). \tag{12.9.1}$$

(Indeed: (12.9.1) implies 12.7.4. To prove this, observe that repeated application of (12.9.1) gives $\varphi(x_{\alpha+\alpha_0} \,|\, \alpha = 1,2,\dots) = \varphi(x_\alpha \,|\, \alpha = 1,2,\dots)$, and the left-hand side expression is obviously independent of x_{α_0}.)

To be precise: Banach's system, as formulated loc. cit. 12.6, consists of conditions amounting to our 12.7.1–12.7.3, (12.9.1) (as stated) together with our 12.8.2, 12.8.3.1. But since the latter are consequences of the former ones (cf. 12.8.2, 12.8.3), we may omit them.

So we see: Our original notion of limits, as well as Banach's notion of limits, obtains by strengthening the system 12.7.1–12.7.4 in two different ways: Ours by replacing 12.7.2 by (12.2.2), and Banach's by replacing 12.7.4 by (12.9.1). We will show:

First: That these two ways of strengthening can never be achieved simultaneously. Second: Yet, when a limit in our sense is given, a Banach limit can be easily constructed with its help. The converse process is probably far less simple.

More precisely:

12.10. *A limit in the sense of* 9.1 *cannot be a Banach limit.*

PROOF. Considering what was said in 12.9 above, this means: Using 12.7.1–12.7.4, condition (12.1.2) contradicts (12.9.1). Indeed: Assume that $\varphi(x_\alpha \,|\, \alpha = 1,2,\dots)$ fulfills all these requirements. Put

$$x_\alpha = 1 \quad \text{for } \alpha \text{ even}, \qquad x_\alpha = 0 \quad \text{for } \alpha \text{ odd}.$$

Then always $x_\alpha = 0, 1, x_\alpha = x_\alpha^2$, so by (12.2.2)

$$\varphi(x_\alpha \,|\, \alpha = 1,2,\dots) = (\varphi(x_\alpha \,|\, \alpha = 1,2,\dots))^2;$$

i.e.

$$\varphi(x_\alpha \,|\, \alpha = 1,2,\dots) = 0, 1. \tag{$*$}$$

On the other hand, always $x_\alpha + x_{\alpha+1} = 1$, so by 12.7.1, 12.7.3 and (12.9.1), $2\varphi(x_\alpha \,|\, \alpha = 1,2,\dots) = 1$; i.e.

$$\varphi(x_\alpha \,|\, \alpha = 1,2,\dots) = 1/2. \tag{$**$}$$

12.11. *Let* $\varphi(x_\alpha \,|\, \alpha = 1, 2, \dots)$ *be a limit in the sense of* 9.1. *Then*

$$\beta(x_\alpha \,|\, \alpha = 1, 2, \dots) = \varphi\left(\frac{x_1 + \cdots + x_\alpha}{\alpha} \,\middle|\, \alpha = 1, 2, \dots\right)$$

is a Banach limit.

PROOF. Observe first that the sequence $(\frac{x_1+\cdots+x_\alpha}{\alpha} | \alpha = 1, 2, \dots)$ is bounded along with $(x_\alpha \,|\, \alpha = 1, 2, \dots)$.

Observe next that the above β fulfills 12.7.1–12.7.3 along with φ. So we must only prove (12.9.1) for β. This amounts to showing $\beta(x_{\alpha+1} - x_\alpha \,|\, \alpha = 1, 2, \dots) = 0$, since β fulfills 12.7.3. Owing to the definition of β this means

$$(*) \qquad \varphi\left(\frac{x_{\alpha-1} - x_1}{\alpha} \,\middle|\, \alpha = 1, 2, \dots\right) = 0.$$

Now since the sequence $(x_\alpha \,|\, \alpha = 1, 2, \dots)$ is bounded, therefore $\lim_{\alpha\to\infty} \frac{x_{\alpha+1} - x_1}{\alpha} = 0$ and hence $(*)$ follows from 12.8.3, since that criterion applies to φ.

12.12. We observed in 12.9 that (12.9.1) is a strengthening of 12.7.4. This will be even better understood by comparing the result which follows, with 12.8.3.

12.12.1. 12.7.1–12.7.3 *imply*: (12.9.1) *is equivalent to this condition*:

12.12.1.1. *If* $\lim_{\beta\to\infty} \frac{x_{\alpha+1}+\cdots+x_{\alpha+\beta}}{\beta}$ (*in the classical sense*) *exists uniformly in* α, *then* $\varphi(x_\alpha \,|\, \alpha = 1, 2, \dots)$ *is equal to it.*

PROOF. Sufficiency of 12.12.1.1: Assume 12.12.1.1. Consider a bounded sequence $(x_\alpha \,|\, \alpha = 1, 2, \dots)$. We want to prove (12.9.1), i.e., owing to 12.7.3 $\varphi(x_{\alpha+1} - x_\alpha \,|\, \alpha = 1, 2, \dots) = 0$. By 12.12.1.1 this relation can be inferred if $\lim_{\beta\to\infty} \frac{x_{\alpha+\beta+1} - x_{\alpha+1}}{\beta} = 0$ uniformly in α. But this is indeed true, owing to the boundedness of $(x_\alpha \,|\, \alpha = 1, 2, \dots)$.

Necessity of 12.12.1.1. Assume (12.9.1). Consider a sequence $(x_\alpha \,|\, \alpha = 1, 2, \dots)$ for which $\lim_{\beta\to\infty} \frac{x_{\alpha+1}+\cdots+x_{\alpha+\beta}}{\beta}$ exists uniformly in α, say

$$\lim_{\beta\to\infty} \frac{x_{\alpha+1} + \cdots + x_{\alpha+\beta}}{\beta} = x^0.$$

Consider an $\varepsilon > 0$. There exists a β_ε such that $\beta > \beta_\varepsilon$ implies $x^0 - \varepsilon \leqq \frac{x_{\alpha+1}+\cdots+x_{\alpha+\beta}}{\beta} \leqq x^0 + \varepsilon$, for all α. Now 12.7.3, (12.9.1) give

$$\varphi\left(\frac{x_{\alpha+1} + \cdots + x_{\alpha+\beta}}{\beta} \,\middle|\, \alpha = 1, 2, \dots\right) = \varphi(x_\alpha \,|\, \alpha = 1, 2, \dots);$$

hence the above inequalities and 12.7.1, 12.8.2 and 12.8.1 give $x^0 - \varepsilon \leqq \varphi(x_\alpha \,|\, \alpha = 1, 2, \dots) \leqq x^0 + \varepsilon$. Since this is true for every $\varepsilon > 0$, therefore $\varphi(x_\alpha \,|\, \alpha = 1, 2, \dots) = x^0$. This completes the proof.

12.13. Comparing what we said at the end of 12.7.2 with the results obtained in 12.8.3.1, 12.11, 12.12.1.1, the following interpretation suggests itself:

Two general "convergence improving" procedures are used in analysis: First, replacing the given sequence by appropriate subsequences. Second, various "summation" methods. These two procedures correspond to two opposite principles: the first to concentrating the attention on ever narrowing parts of the original sequence, the second to spreading it evenly by averaging methods.

Our limit (in the sense of 9.1) is a consequent carrying out of the first principle (cf. the end of 12.7.2), while Banach's procedure combines it with the second one (compare 12.8.3.1 with 12.11, 12.12.1.1).

CHAPTER III

Haar measure

13. Remarks on measures

We return to the considerations and notations of §2. We assume that S is a Hausdorff space, $\lambda(C)$ is a set function defined for all compact sets C, and $\nu(M)$ is the measure generated by $\lambda(C)$ (cf. 4.2). For convenience of reference we give below a set of properties of the space S, and the functions $\lambda(C)$ and $\nu(M)$, and then we establish certain implication relations among them.

13.1.1. $0 \leqq \lambda(C) \leqq \infty$.

13.1.2. $\lambda(C + D) \leqq \lambda(C) + \lambda(D)$.

13.1.3. If $CD = \theta$, $\lambda(C + D) = \lambda(C) + \lambda(D)$.

13.1.4. If $C \subseteqq D$, $\lambda(C) \leqq \lambda(D)$.

13.1.5. If $C^i \neq \theta$, $\lambda(C) > 0$.

13.1.6. $\lambda(C) < \infty$.

13.1.7. $x \to \varphi(x)$ is a homeomorphism of S into itself for which $\lambda(\varphi(C)) = \lambda(C)$ for all C.

13.2. S is locally compact.

13.3.1. $0 \leqq \nu(M) \leqq \infty$.

13.3.2. $\nu(\sum_{i=1}^{\infty} M_i) \leqq \sum_{i=1}^{\infty} \nu(M_i)$.

13.3.3. If $\{M_j\}$ is a sequence of measurable sets (cf. 13.1.1) such that for $k \neq j$, $M_k \cdot M_j = \theta$, then $\nu(\sum_{j=1}^{\infty} M_j) = \sum_{j=1}^{\infty} \nu(M_j)$.

13.3.4. If $M^i \neq \theta$, $\nu(M_i) > 0$.

13.3.5. If $\overline{M}$ is compact, $\nu(M) < \infty$.

13.3.6. $x \to \varphi(x)$ is a homeomorphism of S into itself for which $\nu(\varphi(M)) = \nu(M)$ for all M.

We have already seen that 13.1.1–13.1.3 imply 13.3.1–13.3.3. We shall now proceed to establish the following implication relations, assuming throughout that $\lambda(C)$ satisfies 13.1.1–13.1.3, so that $\nu(M)$ satisfies 13.3.1–13.3.3.

13.4. *The conditions* 13.1.5, 13.2 *together imply* 13.3.4.

13.5. *The conditions* 13.1.4, 13.1.6, 13.2 *together imply* 13.3.5.

13.6. *The condition* 13.1.7 *implies* 13.3.6.

PROOF. For 13.4: Suppose that $M^i \neq \theta$; take $x \in M^i$. Then by 1.7 we may find a compact set C such that $x \in C^i \subseteqq C \subseteqq M^i$, so that (by 13.1.5, 2.3.1, and 2.2.2) $0 < \lambda(C) \leqq \nu(C) \leqq \nu(M)$.

For 13.5: Let O be any open set for which $\overline{O}$ is compact. Then (by 13.1.4, 4.2, and 13.1.6) $\nu(O) \leqq \lambda(\overline{O}) < \infty$, so that 13.3.5 is valid for $M = O$. If M is any set for which $\overline{M}$ is compact, then (by 13.2) we may apply 1.7 and find an open set Q such that $M \subseteqq Q$ and $\overline{Q}$ is compact (our $\overline{M}, S$ are the C, O of 1.7, our Q is the D^i

of 1.7) so that $\nu(M) \leqq \nu(\overline{M}) \leqq \nu(Q) \leqq \nu(\overline{Q}) < \infty$. This proves that 13.3.5 holds in general.

For 13.6: Obvious from (2.1.4) and (2.1.5).

14. Preliminary considerations about groups

14.1. We shall use the following notations. If G is a group we shall denote by ab (for $a, b \in G$) the group theoretical product of a and b, and by a^{-1} (for $a \in G$) the group theoretical inverse of a. If M is any subset of G we write aM (or Ma) for the set of all elements of G of the form ax (or xa) where $x \in M$; we denote by M^{-1} the set of all elements of G of the form x^{-1}, where $x \in M$. If M and N are any two subsets of G, we denote by $M \odot N$ the set of all elements of G of the form xy where $x \in M$, $y \in N$.

If G and H are groups we denote by $G \times H$ their direct product: i.e. $G \times H$ is the set of all pairs (x, y) with $x \in G$, $y \in H$, and $G \times H$ is a group if we define the product of (x', y') by (x'', y'') to be $(x'x'', y'y'')$. Similarly if S and T are topological spaces we denote by $S \times T$ the set of all pairs (x, y) with $x \in S$, $y \in T$; $S \times T$ is a topological space if we define a neighborhood of (x, y) to be the set $O \times P$ where O is a neighborhood (in S) of x and P is a neighborhood (in T) of y. If G and H are topological groups (see below) we denote by $G \times H$ the direct product group with the direct product topology.

Throughout the rest of this chapter we shall assume that *G is a group on which a Hausdorff topology is defined so that G becomes a locally compact topological space and that G is a topological group*: in other words, the functions x^{-1} and xy (with domains G and $G \times G$, respectively, and the range G) are continuous. The principal object of this chapter is to prove that *there exists a measure $\nu(M)$ defined for all Borel sets and satisfying the conditions* 13.3.1–13.3.5 *and which is invariant under left multiplication in G*, i.e. for which $\nu(M) = \nu(aM)$ for every $a \in G$ and every Borel set M. In view of the discussion in §13 it will be sufficient to find a set function $\lambda(C)$ defined for all compact sets C, satisfying 13.1.1–13.1.6, and which is invariant under left multiplication in G. We shall construct such a set function $\lambda(C)$ in the next section (§15); in the remainder we discuss the difficulties to be anticipated and prove some auxiliary results.

14.2. We remark first on the sense in which the distinction between left invariance and right invariance is essential. It is *not* essential in two respects:

14.2.1. If G is a group, then another group G^d obtains from it by replacing the multiplication rule xy (which characterizes G) by yx (which is to characterize G^d). Clearly G^d is the same *set* as G, but a different *group* (except when G is Abelian). Clearly $G^{dd} = G$. While G^d has a different multiplication rule from G, it is easily seen to have the same unit 1 and the same reciprocal x^{-1}. G^d is the *dual* of G.

If G is a topological group, then G^d is obviously one also, with the same topology.

Now by passing from G to G^d the notions of right and left are interchanged. Consider a theorem $\mathcal{T}$ which is true for all groups G, or for all topological groups G which fulfill certain purely topological conditions (e.g.: compactness, or local compactness). Then the theorem which obtains from $\mathcal{T}$ by interchanging in it right and left—its *dual* $\mathcal{T}^d$—is also true, for the same groups: $\mathcal{T}^d$ is proved by applying $\mathcal{T}$ to G^d instead of G.

14.2.2. More specifically: If $\nu(M)$ is any left invariant measure in G, then $\nu'(M) = \nu(M^{-1})$ is a right invariant measure in G, and conversely.

14.2.3. Thus the result announced in 14.1—the existence of a left invariant Lebesgue measure for every locally compact topological group—implies a similar result with respect to right invariance.

This does *not* necessarily imply, however, that any specific left invariant measure $\nu(M)$ (in a particular group G) is also right invariant; nor that any simultaneously left and right invariant measure $\nu(M)$ (for a given group G) must exist. We shall see later (cf. 14.2.4) that there exist groups G for which neither is the case. Groups for which both statements are true are easy to name: Obviously every Abelian group is such, and we shall see that this is also so for all compact groups, as well as certain others.

14.2.4. We conclude this discussion by constructing a left invariant measure $\nu(M)$ in a suitable group G, which is not right invariant. Let G be the group of all linear transformations of the real line, i.e. of the set of all real t, defined by

$$t \to T(t) = xt + y \qquad (x, y \text{ fixed}, x > 0).$$

We make correspond to each $T = T_{x,y}$ the point $z = (x, y)$ of the Euclidean plane: this is a one-to-one correspondence between G and the half-plane $x > 0$. We topologize G with the topology induced by this correspondence. It will be convenient to think of the points z themselves as elements of G. The multiplication law in G is described by the formulae

$$\begin{gathered} z'z'' = (x', y')(x'', y'') = (x'x'', x'y'' + y'), \\ z^{-1} = (x, y)^{-1} = \left(\frac{1}{x}, -\frac{y}{x}\right), \\ 1 = (1, 0). \end{gathered}$$

It is readily verified that G is a topological group.

The measures we construct will be of the form $\nu(E) = \int_E \varphi(z)\, d_m(z)$, where m is planar Lebesgue measure. We observe that if we write, for any $z = (x, y) \in G$,

$$\mu^*(z; E) = m(Ez), \qquad {}^*\mu(z; E) = m(zE),$$

then we have

$$\mu^*(z; E) = xm(E), \qquad {}^*\mu(z; E) = x^2 m(E).$$

(This fact is obvious if E is a rectangle and from this it follows easily for arbitrary E.)

We write

$$\nu(E) = v(\varphi(z); E) = \int_E \varphi(z)\, d_m z.$$

Then we have

$$\begin{aligned} \nu(Ez_0) &= \int_{Ez_0} \varphi(z)\, d_m z = \int_E \varphi(zz_0)\, d\mu^*(z_0; z) = \int_E \varphi(zz_0) x_0\, d_m z, \\ \nu(z_0E) &= \int_{z_0E} \varphi(z)\, d_m z = \int_E \varphi(z_0 z)\, d\mu^*(z_0; z) = \int_E \varphi(z_0 z) x_0^2\, d_m z. \end{aligned}$$

Hence the left and right invariance of $\nu(E)$ is equivalent to

$$\int_E \varphi(z_0 z) x_0^2 \, d_m z = \int_E \varphi(z) \, d_m z \tag{14.2.4.1}$$

and

$$\int_E \varphi(z z_0) x_0 \, d_m z = \int_E \varphi(z) \, d_m z \tag{14.2.4.2}$$

respectively. (14.2.4.1) is satisfied by any φ for which $\varphi(z_0 z)x_0^2 \equiv \varphi(z)$ and (14.2.4.2) by any φ for which $\varphi(z z_0)x_0 \equiv \varphi(z)$ (where, of course, $z_0 = (x_0, y_0)$). Choosing $\varphi(z) = \frac{1}{x^2}$ (or $\varphi(z) = \frac{1}{x}$) yields a measure which is left invariant (or right invariant) but not right invariant (or left invariant).

14.3. The main steps in our construction of a left-invariant set function $\lambda(C)$ (see 14.1) will be the following. Let C and D be compact sets and suppose that $D^i \neq \theta$. Let $n[\frac{C}{D}]$ be the smallest number of left translation copies of D—i.e. of sets of the form aD—that suffice to cover C: the number $n[\frac{C}{D}]$ serves as a comparison of the sizes of C and D. We shall then apply our theory of general limits and form the limit of the numbers $n[\frac{C}{D}]$ (in a sense to be made precise in §15) as D becomes arbitrarily small: the resulting limit measure will serve the purpose of $\lambda(C)$.

The outline of the proof in the preceding paragraph is not quite accurate. In order to illustrate the inaccuracy and motivate our later construction we mention an example. Suppose that G is the Euclidean plane, C is an arbitrary closed, bounded set, and D is the interior and boundary of a circle of radius r. The number $n[\frac{C}{D}]$ is in this case the smallest number $n(r)$ of circles of radius r necessary to cover C. Clearly for all $r > 0$, $n(r)\pi r^2 \geqq m(C)$, where m is planar Lebesgue measure. It has been proved that $\lim n(r)\pi r^2 = \frac{2\pi\sqrt{3}}{9} m(C)$.[1] In other words: Starting with the usual notion of measure, which assigns πr^2 as the area of a circle of radius r, we end up with a different measure, a constant multiple of the original one, when we apply the process outlined above. For this reason we shall not use $n[\frac{C}{D}]$ as a comparison of the sizes of C and D. Rather we shall take a fixed compact set A with $A^i \neq \theta$, and use the quotient $n[\frac{C}{A}]/n[\frac{D}{A}]$ for that comparison. In the sense of the above example it is then to be hoped that the constant factors in numerator and denominator cancel, and that we shall obtain a consistent result.

The use of this quotient $n[\frac{C}{A}]/n[\frac{D}{A}]$ is due to Haar, as well as the most surprising insight that this device suffices to construct a measure. The precise details of our procedure, and also its relationship to Haar's original work, will be given in §15.

We conclude this section by proving a lemma that will be useful later.

14.4. *If A and B are compact subsets of the topological group G, then $A \odot B$ is compact.*

PROOF. The set $A \times B$ is compact in $G \times G$ (by Tychonoff's theorem). The mapping $(x, y) \to xy$ (with domain $G \times G$ and range G) is continuous (since G is a topological group) and maps $A \times B \subseteqq G \times G$ into $A \odot B \subseteqq G$. Hence, by 5.3, $A \odot B$ is compact.

---

[1] See R. Kershner, *The number of circles covering a set*, Amer. J. Math. **61** (1939), 665–671. Observe that $2\pi\sqrt{3}/9 = 1.209\cdots > 1$.

15. The existence of Haar measure

The construction which follows is based, as stated in 14.3, on Haar's pioneer work.[2] Our procedure differs from his in two respects:

First: Haar used the diagonal process (i.e. selection of successive subsequences) to obtain limits (corresponding to 15.4 below). This restricted the applicability of his method to separable groups. We shall apply instead the theory of general limits, as developed in Chapter II. (These two methods are not unrelated, cf. (11.7.2) and (12.13).) A similar procedure has been used by Banach and Saks.

Second: By separating the two notions λ and ν completely (cf. Chapter I and §13) our discussion fits better into the frame of general measure theory.

In connection with the limits mentioned in the first remark above, this too seems worth noting: By use of the left-invariant measure thus constructed and also of the uniqueness theorem (cf. §20), we can prove what amounts essentially to the existence of those limits (in the classical sense, cf. 21.8). But this is only feasible ex post, after the measure has been constructed without knowledge of the existence of these limits.

15.1. We make in this section the assumption mentioned in 14.1: G is a locally compact topological group.

Given a compact set C and an arbitrary set M with $M^i \neq \theta$, we define a set function $n[\frac{C}{M}]$ as follows: Owing to $M^i \neq \theta$, every point of G is contained in a suitable set aM^i. Therefore the sum of all sets aM^i covers C (even G). Now the sets aM^i are all open, along with M^i, and C is compact. Consequently C can be covered by a finite number of them:

$$C \subseteqq a_1 M^i + \cdots + a_n M^i \qquad (n = 0, 1, 2, \ldots). \tag{15.1.1}$$

Denote the smallest n, for which (15.1.1) is possible, by $n[\frac{C}{M}]$.

One verifies immediately:

$$n\left[\frac{C}{M}\right] \leqq n\left[\frac{C}{E}\right] n\left[\frac{E}{M}\right] \tag{15.1.2.1}$$

(here E is subject to both conditions: E compact, $E^i \neq \theta$, while C is only compact.)

$$n\left[\frac{C+D}{M}\right] \leqq n\left[\frac{C}{M}\right] + n\left[\frac{D}{M}\right] \tag{15.1.2.2}$$

(C, D compact). Equally obvious:

$$n\left[\frac{C}{M}\right] \leqq n\left[\frac{D}{M}\right] \quad \text{if } C \subseteqq D, \tag{15.1.2.3}$$

$$n\left[\frac{C}{M}\right] = 0 \quad \text{if and only if } C = \theta, \tag{15.1.2.4}$$

$$n\left[\frac{aC}{M}\right] = n\left[\frac{C}{M}\right]. \tag{15.1.2.5}$$

Consider now a compact set E with $E^i \neq \theta$ (the existence of such a set follows from G's local compactness). The set E will remain fixed throughout the rest of

[2] A. Haar, *Der Massbegriff in der Theorie der kontinuierlichen Gruppen*, Annals of Math., 2d Ser. **34** (1933), 147 169.

this discussion. For any compact set C and any compact set A with $A^i \neq \theta$ (cf. the above remark concerning E) we define the further set function

$$\lambda_A(C) = \frac{n\left[\dfrac{C}{A}\right]}{n\left[\dfrac{E}{A}\right]}. \tag{15.1.3}$$

(The denominator is $\neq 0$ by (15.1.2.3).)

We proceed to derive certain properties of $\lambda_A(C)$:

15.1.4.1. $0 \leqq \lambda_A(C) \leqq \infty$.

15.1.4.2. $\lambda_A(C+D) \leqq \lambda_A(C) + \lambda_A(D)$.

15.1.4.3. $\lambda_A(C+D) = \lambda_A(C) + \lambda_A(D)$ if $(D^{-1} \odot C)(A^{-1} \odot A) = \theta$.

15.1.4.4. $\lambda(C) \leqq \lambda(D)$ if $C \subseteqq D$.

15.1.4.5. $\lambda_A(C) \leqq 1/n[\frac{E}{C}] > 0$ if $C^i \neq \theta$.

15.1.4.6. $\lambda_A(C) \leqq n[\frac{C}{E}] < \infty$.

15.1.4.7. $\lambda_A(aC) = \lambda_A(C)$.

PROOF. For 15.1.4.1: Obvious.

For 15.1.4.2: Immediate by (15.1.3) and (15.1.2.2).

For 15.1.4.3: Immediate by (15.1.3), if

$$n\left[\frac{C+D}{A}\right] = n\left[\frac{C}{A}\right] + n\left[\frac{D}{A}\right] \tag{$*$}$$

is established. This $(*)$ is analogous to (15.1.2.2), and it is obviously true (remembering (15.1.1)) if no aA can have common points with C and D simultaneously. I.e., if $ax = u$, $ay = v$, $x, y \in A$, $u \in C$, $v \in D$ is impossible. Now this would imply $v^{-1}u = y^{-1}x$, and as $v^{-1}u \in D^{-1} \odot C$, $y^{-1}x \in A^{-1} \odot A$, it is indeed excluded by the assumption $(D^{-1} \odot C)(A^{-1} \odot A) = \theta$.

For 15.1.4.4: Immediate by (15.1.3) and (15.1.2.3).

For 15.1.4.5, 15.1.4.6: Immediate by (15.1.3) and (15.1.2.1), by replacing the C, E, M of the latter by our E, C, A and by our C, E, A respectively.

For 15.1.4.7: Immediate by (15.1.3) and (15.1.2.5).

15.2. Before going any further, we wish to point out that 15.1.4.3 is a weakened form of the typical additivity condition, as stated e.g. in (13.1.3). Indeed, the hypothesis of (13.1.3) is

$$C \cdot D = \theta, \tag{15.2.1}$$

while the hypothesis of 15.1.4.3 is

$$(D^{-1} \odot C) \cdot (A^{-1} \odot A) = \theta. \tag{15.2.2}$$

And (15.2.2) implies (15.2.1), since (15.2.1) is equivalent to $1 \notin D^{-1} \odot C$, i.e. to $(D^{-1} \odot C) \cdot (1) = \theta$, while clearly $1 \in A^{-1} \odot A$, i.e. $(A^{-1} \odot A) \supseteqq (1)$.

Therefore (15.2.2) will be nearer to (15.2.1), i.e. 15.1.4.3 to (13.1.3), if the set $A^{-1} \odot A$ is nearer to the set (1). And this is of crucial importance, since the other properties 15.1.4.1–15.1.4.7 of $\lambda_A(C)$ coincide with the remaining requirements of 13.1.1–13.1.7 for $\lambda(C)$.

These considerations motivate the lemma which follows. The lemma says, in effect, that for any two compact sets C, D satisfying (15.2.1) it is possible to find a compact set A with $A^i \neq \theta$, with which they even satisfy (15.2.2). For subsequent application it is convenient to split it into two parts.

15.3.1. *If C, D are compact sets, $C \cdot D = \theta$, then $\widetilde{D^{-1} \odot C}$ is an open set containing* 1.

15.3.2. *If O is an open set containing* 1, *then there exists a compact set A_0 with $1 \in A_0^i$ such that $A_0^{-1} \odot A_0 \subseteqq O$.*

PROOF. For 15.3.1: D^{-1} is compact, since it is a topological image of the compact D. Hence $D^{-1} \odot C$ is compact by 14.4, therefore it is closed by 5.1, and so $\widetilde{D^{-1} \odot C}$ is open. And $C \cdot D = \theta$ implies $1 \notin D^{-1} \odot C$, i.e. $1 \in \widetilde{D^{-1} \odot C}$.

For 15.3.2: Since $y^{-1}x$ is a continuous function of x, y, it is possible to find two open sets O_1, O_2 containing 1, such that $x \in O_1$, $y \in O_2$ imply $y^{-1}x \in O$; i.e., such that $O_2^{-1} \odot O_1 \subseteqq O$. Now apply 1.7 with (1), $O_1 \cdot O_2$ in place of its C, O, and let our A_0 be its D. Then $1 \in A_0^i \subseteqq A_0 \subseteqq O_1 \cdot O_2$, and consequently $A_0^{-1} \odot A_0 \subseteqq O$.

15.4. We proceed to give the details of the limiting process mentioned in 14.3. We shall do this by finding a suitable index set I and an ideal $\mathcal{T}$ of subsets of it, and then we shall refer to our theory of general limits to obtain the existence of a limit function whose independence ideal contains $\mathcal{T}$. The details are as follows.

We will choose for I the family $\mathfrak{J}$ of all compact sets A ($\subseteqq G$) for which $1 \in A^i$. (In other words, $\mathfrak{J}$ is the class of those A for which $\lambda_A(C)$ is defined, except that the requirement $A^i \neq \theta$ has been replaced by the more precise requirement $1 \in A^i$.) Consider a family $\mathfrak{J}' \subseteqq \mathfrak{J}$ and compare the two following properties:

15.4.1.1. There exists an $A_0 \in \mathfrak{J}$ such that $A \in \mathfrak{J}'$ implies $A \nsubseteqq A_O$.

15.4.1.2. There exists an open set O containing 1, such that $A \in \mathfrak{J}'$ implies $A \nsubseteqq O$.

We observe:

15.4.2.1. *The two conditions* 15.4.1.1 *and* 15.4.1.2 *are equivalent. Denote the class of all families $\mathfrak{J}' \subseteqq \mathfrak{J}$ which satisfy these conditions, by $\mathcal{T}$.*

15.4.2.2. *$\mathcal{T}$ is an ideal, $\mathcal{T} \neq$* **I**.

PROOF. For 15.4.2.1: Forward implication: Put $O = A^i$. Inverse implication: Use the local compactness or 1.7.

For 15.4.2.2: $\mathcal{T}$ is an ideal: That $\mathfrak{J}' \in \mathcal{T}$ and $\mathfrak{J}'' \subseteqq \mathfrak{J}'$ imply $\mathfrak{J}'' \in \mathcal{T}$ is clear, and if $\mathfrak{J}', \mathfrak{J}'' \in \mathcal{T}$ then $\mathfrak{J}' + \mathfrak{J}'' \in \mathcal{T}$, because if, e.g., 15.4.1.1 holds for $\mathfrak{J}''$ with A_0' and for $\mathfrak{J}''$ with A_0'', then it holds for $\mathfrak{J}' + \mathfrak{J}''$ with $A_0' \cdot A_0''$.

$\mathcal{T} \neq \mathbf{I}$: Use, e.g., 15.4.1.1. Then $A \nsubseteqq A_0$ excludes that 15.4.1.1 be true for $\mathfrak{J}$ with any A_0; hence $\mathfrak{J} \notin \mathcal{T}$ i.e. $\mathcal{T} \neq \mathbf{I}$.

It would be easy to show that $\mathcal{T} \supseteqq \mathfrak{F}$ (cf. 10.2.2), and that $\mathcal{T}$ is not a prime ideal, but we shall not need these facts.

Both definitions of $\mathcal{T}$ make it clear that the validity of a statement "almost always modulo $\mathcal{T}$" (cf. (11.3)) amounts to its validity for "sufficiently small sets A" in the usual sense.

15.5. Let φ be any limit function defined for all bounded sequences $(x_A \mid A \in I)$ of real numbers x_A such that the independence ideal $\mathfrak{P}$ of φ contains $\mathcal{T}$. (The existence of such a limit function is guaranteed by 9.6.) This is a numerical limit in the sense of §12. Since for every fixed C 15.1.4.1 and 15.1.4.6 imply that $(\lambda_A(C) \mid A \in I)$ is a bounded sequence, we may define

$$\lambda(C) = \varphi(\lambda_A(C) \mid A \in I). \tag{15.5.1}$$

The purpose of this definition may be inferred from the concluding remarks of 15.2 and of 15.4. We are now able to prove that $\lambda(C)$ has the properties discussed in §13.

15.5.2.1. $0 \leqq \lambda(C) < \infty$.

15.5.2.2. $\lambda(C+D) \leqq \lambda(C) + \lambda(D)$.

15.5.2.3. $\lambda(C+D) = \lambda(C) + \lambda(D)$ if $C \cdot D = \theta$.

15.5.2.4. $\lambda(C) \leqq \lambda(D)$ if $C \subseteqq D$.

15.5.2.5. $\lambda(C) > 0$ if $C^i \neq \theta$.

15.5.2.6. $\lambda(aC) = \lambda(C)$.

PROOF. Using the definitory properties of a limit we see:

For 15.5.2.1, 15.5.2.2, 15.5.2.4, 15.5.2.6: Immediate by 15.1.4.1, 15.1.4.2, 15.1.4.4, 15.1.4.7 respectively.

For 15.5.2.5: 15.1.4.5 gives $\lambda(C) \leqq 1/n[\frac{E}{C}] > 0$.

For 15.5.2.3: According to 15.1.4.3.

$$\lambda_A(C+D) = \lambda_A(C) + \lambda_A(D), \tag{$*$}$$

except when

$$(D^{-1} \odot C)\,(A^{-1} \odot A) \neq \theta. \tag{$**$}$$

Now $(**)$ means $A^{-1} \odot A \nsubseteqq \widetilde{D^{-1} \odot C}$. Put $O = \widetilde{D^{-1} \odot C}$ in 15.3.1, and form the A_0 of 15.3.2 for this O. Then $A_0^{-1} \odot A_0 \subseteqq \widetilde{D^{-1} \odot C}$, so $(**)$ implies $A \nsubseteqq A_0$. Thus the set of all A $(\in I)$ for which $(*)$ is not true belongs to $\mathcal{T}$. (Use 15.4.2.1.)

Consequently

$$\begin{aligned}\varphi(\lambda_A(C+D) \mid A \in I) &= \varphi(\lambda_A(C) + \lambda_A(D) \mid A \in I)\\ &= \varphi(\lambda_A(C) \mid A \in I) + \varphi(\lambda_A(D) \mid A \in I)\end{aligned}$$

i.e.

$$\lambda(C+D) = \lambda(C) + \lambda(D).$$

In view of the remark made at the end of 14.1 the existence of a $\lambda(C)$ with the properties 15.5.2.1–15.5.2.6 proves the existence of a left invariant Haar measure.

16. Connection between topology and measure

In §§17, 18, 19 we shall analyze in detail the nature of measure in topological spaces. On the basis of those results, the most important of which is Fubini's theorem (cf. (19.2)), we shall prove the uniqueness of Haar's measure (cf. §20). With the help of these facts, mainly the last-mentioned one, we are going to acquire a much more precise knowledge of the properties of Haar's measure. In the present section we shall merely derive some easily established connections between the behavior of this measure and the group topology.

16.1. Throughout this section we mean by $\nu(M)$ a left invariant Haar measure, i.e. one satisfying $\nu(aM) = \nu(M)$ and 13.3.1–13.3.5, as described in 14.1. (A more elaborate definition and theory will be given, as indicated above, in §§17–18.)

We wish to stress in particular the following point: The theorems 16.2, 16.4 and 16.7.2.1 express the equivalence of certain purely topological properties of G with other properties involving $\nu(M)$. This is to be understood as asserting these equivalences for every fixed choice of $\nu(M)$ (subject to the above requirements for $\nu(M)$). I.e., those properties mean the same thing for every such choice of $\nu(M)$.

We could even replace $\nu(M)$ by a right-invariant Haar measure, since some of those properties do not involve the group properties at all, namely, the topological form of 16.2 and of 16.7.2.1, and 16.4.1 as well as 16.4.3 in 16.4—i.e. we could replace G by G^d (cf. 14.2.1).

16.2. *G is compact if and only if $\nu(S)$ is finite.*

Proof. As stated in 16.1, $\nu(M)$ must be such that the compactness of G implies the finiteness of $\nu(G)$. We need therefore to consider the converse statement only.

Suppose then that G is not compact. Consider a left-invariant Haar measure $\nu(M)$ in G. Let O be any neighborhood of the identity for which $D = \overline{O}$ is compact; then D^{-1} is compact, so that, by 14.4 $C = D \odot D^{-1}$ is also compact. Write $a_1 = 1$; after $a_1, \ldots, a_n \in G$ has been defined, consider the compact set $a_1C + \cdots + a_nC$. If G is not compact we may find $a_{n+1} \in G$ so that $a_{n+1} \notin a_1C + \cdots + a_nC$. In this way we obtain an infinite sequence $\{a_n\}$ of elements of G such that for $p < q$, $a_q \notin a_nC = a_p(D \odot D^{-1})$. This means that a_pD and a_qD are disjoint whenever $p < q$, hence by symmetry whenever $p \neq q$. Therefore

$$\nu(G) \leqq \sum_{i=1}^{\infty} \nu(a_iD) = \infty,$$

as

$$\nu(a_iD) = \nu(D) \geqq \nu(0) > 0.$$

Hence if $\nu(G)$ is finite, G is compact.

16.3. *If G is compact there exists on G a Haar measure $\nu(M)$ for which*

$$\nu(aM) = \nu(M) = \nu(Ma)$$

for all $a \in G$, and $\nu(M) = \nu(M^{-1})$.

Proof. Since G is compact, $G \times G$ is also compact, and there exists in $G \times G$ a left-invariant Haar measure ν^*. For any set $M \subseteqq G$, let $M^* \subseteqq G \times G$ be the set of all $(x, y) \in G \times G$ for which $xy^{-1} \in M$. M^* is the inverse image of M under the continuous mapping $(x, y) \to (xy^{-1})$; hence M^* is a Borel set along with M. We may therefore define a measure in G by $\nu'(M) = \nu^*(M^*)$.

(This is only feasible because the compactness of G, and hence of $G \times G$, implies that $\nu^*(G \times G)$ is finite—and with it every $\nu^*(M^*)$ and every $\nu(M)$. Without this it would be uncertain whether $\nu'(M)$ is ever finite.)

Now

$$(aM)^* = (a, 1)M^*, \qquad (Ma)^* = (1, a^{-1})M^*,$$

so that

$$\nu'(aM) = \nu'(M) = \nu'(Ma).$$

It is immediately verified that $\nu(M) = \nu'(M) + \nu'(M^{-1})$ has all the desired properties.

16.4. In 16.2 we saw that a topological condition (compactness) is equivalent, for groups, to the measure theoretic restriction of finiteness. A similar result in the nonfinite case is the following.

The conditions 16.4.1–16.4.4 (*below*) *are equivalent*:

16.4.1. *G is the sum of countably many compact sets.*

16.4.2. *G is the sum of countably many sets of finite measure.*

16.4.3. *Every collection of pairwise disjoint nonempty open sets is countable.*

16.4.4. *If O is any nonempty open set, there exists a countable sequence $a_1, a_2, \ldots$ in G such that $G = a_1O + a_2O + \cdots$.*

PROOF. 16.4.1→16.4.2: Obvious, since a compact set has finite measure.

16.4.2→16.4.3: Let $\{O_\alpha\}$ be any family of pairwise disjoint nonempty open sets, and let $\{M_n\}$ be a countable sequence of sets of finite measure such that $G = M_1 + M_2 + \cdots$. Consider the set N_{ij} of all α for which $\nu(O_\alpha \cdot M_i) > \frac{1}{j}$. The set N_{ij} is finite (in fact the number of elements of N_{ij} is $\leqq j\nu(M_i)$), so that the set $N = \sum_{i,j} N_{ij}$ is countable. If $\alpha \notin N$, then $\nu(O_\alpha \cdot M_i) = 0$ for $i = 1, 2, \ldots$, so that $\nu(O_\alpha) = 0$, whence $O_\alpha = \theta$. In other words, N exhausts all α's, as was to be proved.

16.4.3→16.4.4: Suppose that there is a nonempty open set O such that for no countable sequence $a_1, a_2, \ldots$ in G do we have $G = a_1O + a_2O + \cdots$.

Consider an $a \in O$. Replacement of O by $a^{-1}O$ does not affect the above properties, and $1 \in a^{-1}O$. So we see: There is no loss in generality if we assume that $1 \in O$. Since xy^{-1} is a continuous function of x, y, it is possible to find two open sets O_1, O_2 containing 1, such that $x \in O_1$, $y \in O_2$ imply $xy^{-1} \in O$; i.e., such that $O_1 \odot O_2^{-1} \subseteqq O$. Put $P = O_1 \odot O_2$. Then P is an open set containing 1, and $P \odot P^{-1} \subseteqq O$.

Now let Ω be the first noncountable ordinal of Cantor. We define a sequence $(a^\alpha \mid \alpha < \Omega)$ in G as follows: Consider an $\alpha < \Omega$. Assume that the a^β for all $\beta < \alpha$ have already been defined. Since $(a^\beta \mid \beta < \alpha)$ is a countable sequence in G, our original hypothesis now necessitates $G \neq \sum_{\beta<\alpha} a^\beta \cdot O$. Choose $a^\alpha \notin \sum_{\beta<\alpha} a^\beta O$ (but, of course, $a^\alpha \in G$).

Thus $\alpha > \beta$ implies $a^\alpha \notin a^\beta O$, and hence $a^\alpha \notin a^\beta(P \odot P^{-1})$; i.e. $a^\alpha P$ and $a^\beta P$ are then disjoint. By symmetry this is true for all $\alpha \neq \beta$. So the $a^\alpha P$, for all $\alpha < \Omega$, form an uncountable collection of pairwise disjoint nonempty open sets. This contradicts 16.4.3. I.e. 16.4.3 implies 16.4.4.

16.4.4→16.4.1. Let O be a neighborhood of the identity for which $\overline{O}$ is compact; then if $\sum_n a_nO$ covers G, so does $\sum_n a_n\overline{O}$.

16.5. The groups G described in 16.2 are *compact*; by analogy we call those described in 16.4 (i.e. those which fulfill the equivalent conditions 16.4.1–16.4.2) ***semi-compact***.

There is an essential difference between compact and semi-compact groups: e.g. the conclusion of 16.3 is not in general true for the latter ones. In fact, in compact groups an alternative method of obtaining the entire theory of measure exists: It is the approach by *almost periodic functions*,[3] which gives existence, uniqueness, and general two-sided invariance of the measure in a simpler way. We do not propose, however, to discuss this procedure here, but will continue with the general method which applies to all (locally compact) groups equally.

On the other hand there is no real difference between the semi-compact case and the general case. Theorem 16.6.3 below shows that the topological structure of any (locally compact) group can be expressed in terms of the structure of a semi-compact one.

16.6. We prove first the following lemma:

16.6.1. *Let G be an arbitrary topological group, G_1 a subgroup of G, and assume $G^i \neq \theta$. Then every set which is the sum of any number of left (right) cosets of G_1 is both open and closed in G. Hence in particular, every left (right) coset of G_1, and thus G_1 itself, is both open and closed in G.*

PROOF. By 14.2.1 it suffices to consider left cosets. Since the complementary set of any sum of left cosets is again a sum of left cosets, and since any set with an open complementary set is necessarily closed, it suffices to establish the openness. Since the sum of any number of open sets is open, it suffices to show that all left cosets of G_1 are open. Since every left coset of G_1 is of the form aG_1, it suffices to prove that G_1 itself is open.

We have $G_1^i \neq \theta$ choose a $b_0 \in G_1^i$. Now consider any $b \in G$. Then b_0, $b \in G_1$, so $bb_0^{-1} \in G_1$. Hence $bb_0^{-1}G_1 = G_1$, and so $bb_0^{-1} \cdot G_1^i = G_1^i$. Consequently $b = bb_0^{-1} \cdot b_0 \in G_1^i$. Since $b \in G_1$ was arbitrary, this means that G_1 is open, and thus completes the proof.

16.6.2. We call a subgroup G_1 of G with $G^i \neq \theta$, i.e. one to which the results of 16.6.1 apply, a *full* subgroup. Clearly 16.6.1 expresses this: A full subgroup G_1 of G embraces the entire topological character of G; all in G that goes beyond G_1, described by the (left or right) coset structure of G_1, is topologically vacuous—since all sums of cosets are identically open and closed. Observe in particular that G_1 is locally compact if G is.

It is in this light that the theorem which follows should be viewed.

16.6.3. *Every (locally compact) topological group G possesses a full subgroup G_1, which is semi-compact (in the sense of 16.5).*

PROOF. Let C be a compact set with $1 \in C^i$. Then define a sequence of sets $C_0, C_1, C_2, \ldots$ as follows:

16.6.3.1. $C_0 = C$.

16.6.3.2. $C_{l+1} = C_l^{-1} \odot C_l$ for $l = 0, 1, 2, \ldots$.

16.6.3.3. $G_1 = C_0 + C_1 + C_2 + \cdots$.

---

[3] J. v. Neumann, *Almost periodic functions in a group*, Trans. Amer. Math. Soc. **36** (1934), 445–492.

We now prove successively:

(α) If C_l is compact, then C_l^{-1} is too, and so is $C_{l+1} = C_l^{-1} \odot C_l$ by 14.4. Now C_0 is compact; consequently all $C_0, C_1, C_2, \ldots$ are compact.

(β) Assume $1 \in C_l^i$. Then $1 \in (C_l^{-1})^i$ as well; hence for every $a \in C_l$, $a \in (C_l^{-1})^i a = (C_l^{-1} a) \subseteqq (C_l^{-1} \odot C_l)^i = C_{l+1}^i$. I.e. $C_l \subseteqq C_{l+1}^i$. We restate:

$$1 \in C_l^i \quad \text{implies} \quad C_l \subseteqq C_{l+1}^i.$$

(γ) The final result of (β) shows that $1 \in C_l^i$ implies $C_l \subseteqq C_{l+1}^i$; hence $C_l^i \subseteqq C_{l+1}^i$; hence $1 \in C_{l+1}^i$. Now $1 \in C_0^i$; hence $1 \in C_l^i$ for all $l = 0, 1, 2, \ldots$. Therefore (β) permits us to assert unconditionally:

Always $C_l \subseteqq C_{l+1}^i$.

(δ) Now (γ) implies $C_l \subseteqq C_{l+1}$, i.e.

$$C_0 \subseteqq C_1 \subseteqq C_2 \subseteqq \cdots .$$

(ε) 16.6.3.2 and (δ) give, considering 16.6.3.3, that $G_1 \supseteqq G_1^{-1} \odot G_1$. This means obviously that G_1 is a subgroup of G.

(ϑ) Since $1 \in C_0^i$, hence $1 \in G_1^i$, we have $G_1^i \neq \theta$. I.e. G_1 is a full subgroup of G.

(ξ) Owing to (α) G_1 is semi-compact (cf. 16.5).

The proof is thus completed

16.6.4. The semi-compact full subgroup G_1 of G, which we discussed above, is in general not a uniquely determined object. We can choose $G_1 = G$ if and only if G itself is semi-compact.

For a general G, we have, however, this:

Given a sequence $D_0, D_1, D_2, \ldots$ of compact subsets of G, the semi-compact full subgroup G_1 of G (cf. 16.6.3) can be chosen so that $G_1 \supseteqq D_l$ for all $l = 0, 1, 2, \ldots$.

PROOF. Repeat the construction of the proof of 16.6.3, with the same 16.6.3.1 and 16.6.3.3, only replacing 16.6.3.2 by

$$C_{l+1} = C_l^{-1} \odot C_l + D_l \quad \text{for } l = 0, 1, 2, \ldots . \tag{16.6.4.1}$$

Then the considerations of (α)–(ξ) in the proof of 16.6.3 are absolutely unaffected; hence G_1 is again a semi-compact full subgroup of G. And furthermore $G_1 \supseteqq C_{l+1} \supseteqq D_l$ for all $l = 0, 1, 2, \ldots$, as desired.

16.7. We conclude this section by discussing another topological property which is equivalent to a property of the measure $\nu(M)$. (Cf. 16.1.)

A topological space S is *discrete* if every $p \in S$ has itself—i.e. the one element set (p)—as a neighborhood. Then obviously every set $M \subseteqq S$ is both open and closed, and M is compact if and only if it is finite. (Proof of the last statement: Every finite set M is obviously compact. Conversely, let M be compact. $M = \sum_{p \in M}(p)$, and here each set (p) is open. Hence $M \subseteqq (p_1) + \cdots + (p_n)$; i.e. M is finite.)

The lemma which follows is of some interest.

16.7.1. *A locally compact S is discrete if and only if every compact set $C \subseteqq S$ is finite.*

PROOF. Necessity: Established above.

Sufficiency: Assume that S is locally compact and that every compact $C \subseteqq S$ is finite. Consider a $p \in S$. p possesses a neighborhood O with compact closure $\overline{O}$.

Thus $\overline{O}$, and with it O, is finite. Denote the elements $\neq p$ of O by $p_1, \ldots, p_m$. For every $i = 1, \ldots, m$ form a neighborhood P_i of p with $p_i \notin P_i$. Then $Q = O \cdot P_1 \cdot \cdots \cdot P_m$ is also a neighborhood of p. Hence $p \in Q$, never $p_i \in Q$ since never $p_i \in P_i$, and Q has no other elements than $p, p_1, \ldots, p_m$, since $Q \subseteqq O$. Thus $P = (p)$. This proves the discreteness of S.

We can now reach our real objective:

16.7.2.1. *G is not discrete if and only if $\nu((a)) = 0$ for every one-element set (a).*

16.7.2.2. *If G is discrete, then $\nu(M) = Cn(M)$, where C is a constant $> 0, < \infty$, and $n(M)$ is the number of elements of M.*

PROOF. For 16.7.2.1: Sufficiency: Assume that always $\nu((a)) = 0$. Since $(a) \neq \theta$, this excludes the openness of (a) by 13.3.4. So G cannot be discrete.

Necessity: Assume that $\nu((a_0)) \neq 0$ for an appropriate $a_0 \in G$. Put $\nu((a_0)) = C$. Then $C > 0$ by assumption, and $C < \infty$ because the set (a_0) is compact. Owing to the left invariance $\nu((a)) = \nu((a_0))$ for every $a \in G$, since $(a) = aa_0^{-1} \cdot (a_0)$. So we see:

(*) $\nu((a)) = C$ for all $a \in G$, where C is a constant $> 0, < \infty$.

Consequently

(**) $\nu(M) = Cn(M)$ for every finite set $M \subseteqq G$.

Now for every compact set D, $\nu(D)$ is finite; hence by (**), which holds for all sets M (cf. the proof of 16.7.2.2 below), $n(D)$ is finite; i.e. D is a finite set. Therefore G is discrete by 16.7.1.

For 16.7.2.2: The desired formula was established for all finite sets M under (**) above. If M is infinite, then we can find an n-element set $M_n \subseteqq M$ for every $n = 1, 2, \ldots$. Now $\nu(M) \geqq \nu(M_n) = C_n$ for all $n = 1, 2, \ldots$; i.e. $\nu(M) = \infty$. Since $n(M) = \infty$ also, this proves (**) for all infinite sets M too.

CHAPTER IV

Uniqueness

17. Set theory

In this section we shall discuss certain combinatorial properties of classes of sets which we shall have occasion to apply later. Let S be an arbitrary set: we define five types of classes of subsets of S.

17.1. A class $\mathcal{R}$ of sets is a *ring* if $A, B \in \mathcal{R}$ implies $A + B$, $AB \in \mathcal{R}$.

17.2. A class of sets is of *type* $\mathcal{D}$ if, whenever it contains A and B, it also contains a finite number of pairwise disjoint sets, $C_1, \ldots, C_n$, such that

$$\widetilde{AB} = C_1 + \cdots + C_n.$$

We shall use the letter $\mathcal{D}$ to denote the generic class of this type.

17.3. A class $\mathcal{F}$ of sets is a *field* if $A, B \in \mathcal{F}$ implies $A + B$, $\widetilde{AB} \in \mathcal{F}$. We observe that if $A, B \in \mathcal{F}$, then $AB = (A + B)\widetilde{(A' + B')}$ where $A' = (A + B)\widetilde{A}$, and $B' = (A + B)\widetilde{B}$, so that $AB \in \mathcal{F}$.

17.4. A class $\mathcal{B}$ of sets is a *Borel field* if $A_1, A_2, \ldots \in \mathcal{B}$ implies $A_1 + A_2 + \cdots \in \mathcal{B}$ and $A, B \in \mathcal{B}$ implies $A\widetilde{B} \in \mathcal{B}$. As above we note that since

$$A_1 A_2 \cdots = (A_1 + A_2 + \cdots)\widetilde{(A_1' + A_2' + \cdots)},$$

where $A_i' = (A_1 + A_2 + \cdots)\widetilde{A}_i$, it follows that $A_1 A_2 \cdots \in \mathcal{B}$.

17.5. A class $\mathcal{M}$ of sets is a *monotone Borel ring* if $A_1 \leqq A_2 \leqq \cdots$ and $A_1 \in \mathcal{M}$ for $i = 1, 2, \ldots$ implies $A_1 + A_2 + \cdots \in \mathcal{M}$ and at the same time $A_1 \geqq A_2 \geqq \cdots$ and $A_i \in \mathcal{M}$ for $i = 1, 2, \ldots$ implies $A_1 \cdot A_2 \cdots \in \mathcal{M}$.

We note that the set of all subsets of S fulfills the conditions of all five definitions, and that the intersection of any number of rings, fields, Borel fields, or monotone Borel rings, respectively, is again a class of the same type. Hence with an arbitrary class $\mathcal{C}$ of sets we may associate the intersection of all rings containing $\mathcal{C}$: this intersection is itself a ring containing C and is the smallest such ring. We denote it by $\mathcal{R}(\mathcal{C})$. Similarly we denote by $\mathcal{F}(\mathcal{C}), \mathcal{B}(\mathcal{C})$ and $\mathcal{M}(\mathcal{C})$, respectively, the smallest field, Borel field, or monotone Borel ring containing $\mathcal{C}$. The above statements do not apply, however, to sets of type $\mathcal{D}$: an intersection of even two such sets need not be of type $\mathcal{D}$ and one cannot define the smallest set of type $\mathcal{D}$ containing C at all. We shall instead use $\mathcal{D}'(\mathcal{C})$ to denote the collection of all sets of the form $C_1 + \cdots + C_n$, where $C_i \in \mathcal{C}$ and for $i \neq j$, $C_i C_j = \theta$, $i, j = 1, \ldots, n$.

17.6. *Let $\mathcal{R}$ be a ring, $\mathfrak{A}$ the class of all sets of the form $A\widetilde{B}$, where A and B are arbitrary elements of $\mathcal{R}$ and $\mathfrak{A}'$ the class of all sets of the form $A\widetilde{B}$ where A and B are in $\mathcal{R}$ and $A \geqq B$. Then $\mathfrak{A}$ is of type $\mathcal{D}$ and $\mathfrak{A} = \mathfrak{A}'$.*

PROOF. Clearly $\mathfrak{A}' \leqq \mathfrak{A}$; on the other hand, $A, B \in \mathcal{R}$ implies $A\widetilde{B} = A(\widetilde{AB}) \in \mathfrak{A}'$ so that $\mathfrak{A} = \mathfrak{A}'$. Consider now the difference of any two sets in $\mathfrak{A}'$: i.e. suppose that $A_1, B_1, A_2, B_2 \in \mathcal{R}$, $A_1 \geqq B_1$, $A_2 \geqq B_2$, and consider $(A_1\widetilde{B}_1)(\widetilde{A_2\widetilde{B}_2})$. We have

$$(A_1\widetilde{B}_1)(\widetilde{A_2\widetilde{B}_2}) = (A_1\widetilde{B}_1)(\widetilde{A}_2 + B_2) = A_1\widetilde{B}_1\widetilde{A}_2 + A_2\widetilde{B}_1B_2$$

$$= A_1(\widetilde{A_2 + B_1}) + (A_1B_2)\widetilde{B}_1.$$

Since $(\widetilde{A_2 + B_1}) \leqq \widetilde{A}_2 \leqq \widetilde{B}_2$, the terms of the last written sum are disjoint, and since $\mathcal{R}$ is a ring, each addend is in $\mathfrak{A}$. This completes the proof that $\mathfrak{A}$ is of type $\mathcal{D}$.

17.7. *If $\mathfrak{A}$ is a class of type $\mathcal{D}$, then $\mathcal{F}(\mathfrak{A}) = \mathcal{D}'(\mathfrak{A})$.*

PROOF. It is clear that $\mathcal{D} = \mathcal{D}'(\mathfrak{A}) \leqq \mathcal{F} \leqq \mathcal{F}(\mathfrak{A})$; we shall prove that $\mathcal{D} = \mathcal{F}$ by showing that $\mathcal{D}$ is a field.

17.7.1. It is clear from the definition of $\mathcal{D}(\mathfrak{A})$ that $D_i \in \mathcal{D}$ for $i = 1, \ldots, n$; $D_iD_j = \theta$ for $i \neq j$ implies $D_1 + \cdots + D_n \in \mathcal{D}$.

17.7.2. If $A, B \in \mathfrak{A}$, the $A\widetilde{B} \in \mathcal{D}$.

17.7.3. If $A \in \mathcal{D}$ and $B \in \mathfrak{A}$, then $A\widetilde{B} \in \mathcal{D}$. For by hypothesis we may write A as a disjoint sum of sets of $\mathfrak{A}$, $A = A_1 + A_2 + \cdots + A_n$, so that

$$A\widetilde{B} = (A_1 + \cdots + A_n)\widetilde{B} = A_1\widetilde{B} + \cdots + A_n\widetilde{B}.$$

By 17.7.2, $A_i\widetilde{B} \in \mathcal{D}$ and it follows from 17.7.1 that $A\widetilde{B} \in \mathcal{D}$.

17.7.4. If $A, B \in \mathcal{D}$, then $A\widetilde{B} \in \mathcal{D}$. For if $B = B_1 + \cdots + B_m$, where the B_i are pairwise disjoint sets of $\mathfrak{A}$, then

$$A\widetilde{B} = A(\widetilde{B_1 + \cdots + B_m}) = A\widetilde{B}_1\widetilde{B}_2\cdots\widetilde{B}_m,$$

and the desired result follows by repeated application of 17.7.3.

17.7.5. If $A, B \in \mathcal{D}$, $A + B \in \mathcal{D}$. For we have $A + B = A\widetilde{B} + B$. The latter sum has disjoint addends which (by 17.7.4) belong to $\mathcal{F}_{\mathcal{D}}$; hence (by 17.7.1) it belongs to $\mathcal{D}$.

Together the statements 17.7.4 and 17.7.5 merely assert that is a field, as was to be proved.

17.8. *If $\mathcal{F}$ is a field, then $\mathcal{B}(\mathcal{F}) = \mathcal{M}(\mathcal{F})$.*

PROOF. The structure of this proof is similar to the one given above. We observe that $\mathcal{M} = \mathcal{M}(\mathcal{F}) \leqq \mathcal{B} = \mathcal{B}(\mathcal{F})$ and we shall complete the proof by showing that $\mathcal{M}$ is a Borel field. We remark that it is sufficient to prove that $\mathcal{M}$ is a field. For if $\mathcal{M}$ is a field and $A_i \in \mathcal{M}$, $i = 1, 2, \ldots$, then $A_i' = A_1 + \cdots + A_i \in \mathcal{M}$, whence (since $\mathcal{M}$ is, by definition, a monotone Borel ring)

$$\sum_{i=1}^{\infty} A_i = \sum_{i=1}^{\infty} A_i' \in \mathcal{M}.$$

To prove that $\mathcal{M}$ is a field we introduce three auxiliary classes of sets as follows:

Let

$$\begin{Bmatrix} \mathcal{M}' \\ \mathcal{M}'' \\ \mathcal{M}''' \end{Bmatrix}$$

be the class of all sets A such that for all

$$B \in \begin{Bmatrix} \mathcal{F} \\ \mathcal{M}' \\ \mathcal{M} \end{Bmatrix}$$

we have $A+B, A\widetilde{B}, \widetilde{A}B \in \mathcal{M}$. It is easily verified (cf. 17.5) that $\mathcal{M}', \mathcal{M}''$ and $\mathcal{M}'''$ are monotone Borel rings since $\mathcal{M}$ is one. If $A \in \mathcal{M}'$ and $B \in \mathcal{F}$, then (by the definition of $\mathcal{M}'$) $A+B, A\widetilde{B}, \widetilde{A}B \in \mathcal{M}$ so that (interchanging the roles of A and B and using the definition of $\mathcal{M}''$), $B \in \mathcal{M}''$. This means that $\mathcal{F} \leqq \mathcal{M}''$ and therefore

$$\mathcal{M} = \mathcal{M}(\mathcal{F}) \leqq \mathcal{M}''. \tag{17.8.1}$$

Hence if $A \in \mathcal{M}$ (so that $A \in \mathcal{M}''$) and $B \in \mathcal{M}'$, then $A+B, A\widetilde{B}, \widetilde{A}B \in \mathcal{M}$ and (again interchanging the roles of A and B and using the definition of $\mathcal{M}'''$) it follows that $B \in \mathcal{M}'''$ or,

$$\mathcal{M}' \leqq \mathcal{M}'''. \tag{17.8.2}$$

Finally, the fact that $\mathcal{F}$ is a field implies that $\mathcal{F} \leqq \mathcal{M}'$ so that

$$\mathcal{M} = \mathcal{M}(\mathcal{F}) \leqq \mathcal{M}'. \tag{17.8.3}$$

Combining (17.8.2) and (17.8.3) we obtain

$$\mathcal{M} \leq \mathcal{M}'''$$

and this implies that for $A, B \in \mathcal{M}$ (and therefore $A \in \mathcal{M}'''$, $B \in \mathcal{M}$) we have $A+B, A\widetilde{B}, \widetilde{A}B \in \mathcal{M}$; i.e. $\mathcal{M}$ is a field.

17.9. If $\mathcal{E}$ is any class of sets and A_0 is any set, we denote by $\mathcal{E}^{A_0}$ the class of all sets of the form AA_0, where $A \in \mathcal{E}$.

If $\mathcal{R}$ is a ring and $A_0 \in \mathcal{R}$, then

17.9.1. *$\mathcal{R}^{A_0}$ is the set of all $B \in \mathcal{R}$ with $B \leqq A_0$, and*

17.9.2. $\mathcal{B}(\mathcal{R}^{A_0}) = (\mathcal{B}(\mathcal{R}))^{A_0}$.

Proof. For 17.9.1: Every element B of $\mathcal{R}^{A_0}$ has the form $B = AA_0$, $A \in \mathcal{R}$; hence $B \in \mathcal{R}, B \leqq A_0$. Conversely, $B \in \mathcal{R}$, $B \leqq A_0$ implies $B = BA_0 \in \mathcal{R}^{A_0}$.

For 17.9.2: By 17.9.1 $\mathcal{R}^{A_0} \leqq \mathcal{R}$ so that $\mathcal{B}(\mathcal{R}^{A_0}) \leq \mathcal{B}(\mathcal{R})$. Since the sets $B \leqq A_0$ form a Borel field which contains all $\mathcal{R}^{A_0}$ and therefore $\mathcal{B}(\mathcal{R}^{A_0}), \mathcal{B}(\mathcal{R}^{A_0}) \leqq (\mathcal{B}(\mathcal{R}))^{A_0}$. Conversely, the sets A with $AA_0 \in \mathcal{B}(\mathcal{R}^{A_0})$ form a Borel field. Since $A \in \mathcal{R}$ implies $AA_0 \in \mathcal{R}^{A_0} \leqq \mathcal{B}(\mathcal{R}^{A_0})$ this Borel field contains $\mathcal{R}$ and therefore $\mathcal{B}(\mathcal{R})$. Hence $(\mathcal{B}(\mathcal{R}))^{A_0} \leqq \mathcal{B}(\mathcal{R}^{A_0})$ as was to be proved.

17.10. *If $\mathcal{R}$ is a ring and $A \in \mathcal{B}(\mathcal{R})$, then there exists a sequence $A_1, A_2, \dots$ of sets in $\mathcal{R}$ such that*

$$A_1 \leqq A_2 \leqq A_3 \leqq \cdots, \tag{17.10.1}$$

$$A = AA_1 + AA_2 + \cdots. \tag{17.10.2}$$

PROOF. The class of all subsets of all sums of the form $A_1 + A_2 + \cdots$ with $A_i \in \mathcal{R}$ is a Borel field containing $\mathcal{R}$ and therefore containing $\mathcal{B}(\mathcal{R})$. This proves that every $A \in \mathcal{B}(\mathcal{R})$ can be written in the form (17.10.2) with $A_i \in \mathcal{R}$; replacing A_i by $A_1 + \cdots + A_i$ we obtain (17.10.1).

Combining the results in 17.6–17.10 we may sum up as follows:

17.11. *Let $\mathcal{R}$ be any ring of subsets of S. We obtain the smallest Borel field $\mathcal{B}(\mathcal{R})$ containing $\mathcal{R}$ by the following sequence of steps.*

(0) *For any $A_0 \in \mathcal{R}$ form the class $\mathcal{R}^{A_0}$ of all sets $A \in \mathcal{R}$ with $A \leqq A_0$.*

(I) *Form the class $\mathcal{D}_{A_0}$ of all sets $A\widetilde{B}$, with $A, B \in \mathcal{R}^{A_0}$, $A \leqq B$.*

(II) *Form the class $\mathcal{F}_{A_0}$ of all sets $A_1 + \cdots + A_n$, where the A_i are pairwise disjoint sets of $\mathcal{D}_{A_0}$.*

(III) *Form the monotone Borel ring $\mathcal{M}_{A_0} = \mathcal{M}(\mathcal{F}_{A_0})$.*

(IV) *Form the class $\mathcal{B}$ of all sets $A = A_1 + A_2 + \cdots$, where $A_1 \leqq A_2 \leqq \cdots$ and $A_i \in \mathcal{F}_{A_0}$ for some $A_0 \in \mathcal{R}$, $i = 1, 2, \dots$.*

Then $\mathcal{B}(\mathcal{R}) = \mathcal{B}$.

We observe that if $\mu(A)$ is a completely additive nonnegative measure defined on $\mathcal{B} = \mathcal{B}(\mathcal{R})$ which is finite for $A \in \mathcal{R}$, then each of the steps (0)–(IV) determines the measure of the sets formed in terms of the sets given.

Indeed (0) is only a selection from $\mathcal{R}$; (I) uses the operation $A\widetilde{B}$ with $A \geqq B$ and $\mu(A), \mu(B)$ finite (since $A, B \leqq A_0$), so that $\mu(A\widetilde{B}) = \mu(A) = \mu(B)$; (II) uses the operation $A_1 + \cdots + A_n$ where the A_i are pairwise disjoint sets, so that $\mu(A_1 + \cdots + A_n) = \mu(A_1) + \cdots + \mu(A_n)$. (III) is based on the operations $A_1 + A_2 + \cdots$ with $A_1 \leqq A_2 \leqq \cdots$ and $A_1A_2\cdots$ with $A_1 \geqq A_2 \geqq \cdots$. In the latter case all $\mu(A_i)$ are finite (since $A_i \leqq A_0$), so that $\mu(A_1A_2\cdots)$ as well as $\mu(A_1 + A_2 + \cdots)$ equals $\lim_{i\to\infty} \mu(A_i)$. (IV) again uses the operation $A_1 + A_2 + \cdots$ with $A_1 \leqq A_2 \leqq \cdots$.

Thus many properties of μ on $\mathcal{B}(\mathcal{R})$ can be established by proving that they are valid for μ on $\mathcal{R}$. By way of illustration we mention that if two measures μ and ν are equal on $\mathcal{R}$, then they are equal on $\mathcal{B}$. We shall make many applications of this type of 17.11.

An example, often used in the sequel, of a ring of sets is given by the class of all compact sets in any topological space.

18. Regularity

In order to prove the uniqueness of Haar measure we have to show that among all measures there is essentially only one that satisfies the conditions satisfied, in particular, by the measure we derived in the existence proof. In order to make this statement more specific we have to specify exactly what we are going to mean by a measure, and what the characteristic properties are. Throughout the remainder of our work when we discuss measures on locally compact topological spaces we make the following assumptions. $\mathcal{R}$ is the ring of all compact sets; $\mathcal{B} = \mathcal{B}(\mathcal{R})$ is the smallest Borel field containing $\mathcal{R}$. Sets of $\mathcal{B}$ we shall call Borel sets and we reserve, from now on, this terminology for sets of $\mathcal{B}$ exclusively. (Thus our terminology is here at variance with the usual one in topological spaces—the one we used in Chapter I—according to which the Borel sets are the members of the smallest Borel field containing all open sets.) A measure is a completely additive and nonnegative set function with domain $\mathcal{B}$ which is finite for compact sets, and which satisfies another regularity condition, which will be discussed in this section. (Cf. also 4.2.)

Consider the following two conditions on a measure ν:

18.1. For all M, $\nu(M) = \sup \nu(C)$, $C \leqq M$, C compact.

18.2. For all M, $\nu(M) = \inf \nu(O)$, $M \leqq O$, O open.

(We write the symbols $\nu(M), \nu(O)$, etc., with the understanding that the arguments M, O, etc., are Borel sets. Observe that, according to (2.1.5), 2.3.2 and 4.3, the measure given by our existence proof (§15) satisfies both these conditions.)

18.3. *Both conditions* 18.1 *and* 18.2 *are hereditary under steps* (II), (III), *and* (IV) *of* 17.11: *i.e.* (*in the notation of* 17.11) *their validity for all sets M in all $\mathcal{D}_{A_0}$ guarantees their validity for all $M \in \mathcal{B} = \mathcal{B}(\mathcal{R})$, where $\mathcal{R}$ is the ring of all compact sets.*

Proof. We conclude from the discussion following 17.11 that we must prove 18.1 and 18.2 to be hereditary under these three operations:

(i) $A_1 + \cdots + A_n$, where the A_i are pairwise disjoint.

(ii) $A_1 + A_2 + \cdots$, where $A_1 \leqq A_2 \leqq \cdots$.

(iii) $A_1 \cdot A_2 \cdot \cdots$, where $A_1 \geqq A_2 \geqq \cdots$.

Throughout (i)–(iii) every $A_i \leqq A_{0,i}$ for some $A_{0,i} \in \mathcal{R}$; hence all $\nu(A_i)$ are finite.

Thus six assertions result, which we now proceed to establish.

For 18.1 and (i): Choose any $\varepsilon > 0$. Choose $C_i \leqq A_i$ with $\nu(C_i) \geqq \nu(A_i) - \varepsilon/n$. Put $C = C_1 + \cdots + C_n$. Clearly $C \leqq A_1 + \cdots + A_n$. Since the C_i are pairwise disjoint, along with the A_i, so

$$\begin{aligned}\nu(C) = \nu(C_1) + \cdots + \nu(C_n) &\geqq \nu(A_1) + \cdots + \nu(A_n) - \varepsilon \\ &= \nu(A_1 + \cdots + A_n) - \varepsilon.\end{aligned}$$

For 18.1 and (ii): Choose any α with $\nu(A_1 + A_2 + \cdots) > \alpha$. This means $\lim_{i\to\infty} \nu(A_i) > \alpha$ so we can choose a j with $\nu(A_j) > \alpha$. Now choose $C \leqq A_j$ with $\nu(C) > \alpha$. Clearly $C \leqq A_1 + A_2 + \cdots$.

For 18.1 and (iii): Choose any $\varepsilon > 0$. Choose $C_i \leqq A_i$ with $\nu(C_i) \geqq \nu(A_i) - \varepsilon/2^i$. Put $C = C_1 \cdot C_2 \cdot \cdots$. Clearly $C \leqq A_1 \cdot A_1 \cdot \cdots$. Owing to the easily verifiable set-theoretical identity,

$$(A_1 \cdot A_2 \cdot \cdots) \cdot \widetilde{(C_1 \cdot C_2 \cdot \cdots)} \leqq (A_1 \cdot \widetilde{C}_1) + (A_2 \cdot \widetilde{C}_2) + \cdots$$

we have

$$\nu(A_1 \cdot A_2 \cdot \cdots) - \nu(C) = \nu(A_1 \cdot A_2 \cdot \cdots) - \nu(C_1 \cdot C_2 \cdot \cdots)$$

$$= \nu((A_1 \cdot A_2 \cdot \cdots) \cdot \widetilde{(C_1 \cdot C_2 \cdot \cdots)}) \leqq \nu(A_1 \cdot \widetilde{C}_1) + \nu(A_2 \cdot \widetilde{C}_2) + \cdots$$

$$= (\nu(A_1) - \nu(C_1)) + (\nu(A_2) - \nu(C_2)) + \cdots \leqq \varepsilon$$

i.e. $\nu(C) \geqq \nu(A_1 \cdot A_2 \cdot \cdots) - \varepsilon$.

For 18.2 and (i): Choose any $\varepsilon > 0$. Choose $O_i \geqq A_i$ with $\nu(O_i) \leqq \nu(A_i) + \varepsilon/n$. Put $O = O_1 + \cdots + O_n$. Clearly $O \geqq A_1 + \cdots + A_n$. Further

$$\nu(O) \leqq \nu(O_1) + \cdots + \nu(O_n) \leqq \nu(A_1) + \cdots + \nu(A_n) + \varepsilon$$

$$= \nu(A_1 + \cdots + A_n) + \varepsilon.$$

For 18.2 and (ii): Choose any $\varepsilon > 0$. Choose $O_i \geqq A_i$ with $\nu(O_i) \leqq \nu(A_i) + \varepsilon/2^i$. Clearly $O \geqq A_1 + A_2 + \cdots$. Owing to the easily verifiable set-theoretical identity

$$(O_1 + O_2 + \cdots) \cdot \widetilde{(A_1 + A_2 + \cdots)} \leqq (O_1 \cdot \widetilde{A}_1) + (O_2 \cdot \widetilde{A}_2) + \cdots$$

we have

$$\nu(O) - \nu(A_1 + A_2 + \cdots) = \nu(O_1 + O_2 + \cdots) - \nu(A_1 + A_2 + \cdots)$$

$$= \nu((O_1 + O_2 + \cdots) \cdot \widetilde{(A_1 + A_2 + \cdots)}) \leqq \nu(O_1 \cdot \widetilde{A}_1) + \nu(O_2 \cdot \widetilde{A}_2) + \cdots$$

$$= (\nu(O_1) - \nu(A_1)) + (\nu(O_2) - \nu(A_2)) + \cdots \leqq \varepsilon;$$

i.e. $\nu(O) \leqq \nu(A_1 + A_2 + \cdots) + \varepsilon$.

For 18.2 and (iii): Choose any α with $\nu(A_1 \cdot A_2 \cdot \cdots) < \alpha$. This means $\lim_{i\to\infty} \nu(A_i) < \alpha$, so we can choose a j with $\nu(A_j) < \alpha$. Now choose $O \geqq A_j$ with $\nu(O) < \alpha$. Clearly $O \geqq A_1 \cdot A_2 \cdot \cdots$.

18.4. We show next that for the general validity of properties 18.1 and 18.2 we need require their validity for only a comparatively small class of sets. In the following proofs we make use of a local compactness of the space.

18.4.1. *The validity of* 18.1 *for all open sets M with compact closure implies its validity for all Borel sets.*

18.4.2. *The validity of* 18.2 *for all compact sets implies its validity for all Borel sets.*

Proof. We observe first that, due to 18.3, in both cases we have only to prove that the conditions are valid for all sets of the form $C\widetilde{D}$ where C and D are compact sets, $D \leqq C$. We suppose accordingly that we are given two such compact sets.

For 18.4.1: Choose any $\varepsilon > 0$, and an open set $O \geqq C$ with compact closure. Then $O\widetilde{D}$ is also an open set with compact closure so that, by assumption, we can find a compact set $C' \leqq O\widetilde{D}$ with $\nu(C') \geqq \nu(O\widetilde{D}) - \varepsilon$. Thus

$$\nu(O\widetilde{D}\widetilde{C'}) = \nu(O\widetilde{D}) - \nu(C') \leqq \varepsilon.$$

Since $C \cdot C'$ is compact and $C \cdot C' = C \cdot O \cdot \widetilde{D} = C \cdot \widetilde{D}$, and since $(C\widetilde{D}) \cdot (\widetilde{C \cdot C'}) = C \cdot (\widetilde{D} \cdot \widetilde{C'}) \leqq O \cdot \widetilde{D} \cdot \widetilde{C'}$, we have

$$\nu(C\widetilde{D}) - \nu(C \cdot C') = \nu((C\widetilde{D}) \cdot (\widetilde{C \cdot C'})) \leqq \nu(O \cdot \widetilde{D} \cdot \widetilde{C'}) \leqq \varepsilon;$$

i.e. $\nu(CC') \geqq \nu(C \cdot D') - \varepsilon$.

For 18.4.2: Choose any $\varepsilon > 0$. By assumption we can find an open set $O \geqq C$ with $\nu(O) \leqq \nu(C) + \varepsilon$. Thus $\nu(O \cdot \widetilde{C}) = \nu(O) - \nu(C) \leq \varepsilon$. Now $O \cdot \widetilde{D}$ is open and $O \cdot \widetilde{D} \geqq C \cdot \widetilde{D}$, and since

$$O\widetilde{D} \cdot (\widetilde{C \cdot \widetilde{D}}) = (O \cdot \widetilde{C}) \cdot \widetilde{D} \leqq O \cdot \widetilde{C},$$

we have

$$\nu(O \cdot \widetilde{D}) - \nu(C\widetilde{D}) = \nu((O \cdot \widetilde{D}) \cdot (\widetilde{C \cdot \widetilde{D}})) \leqq \nu(O \cdot \widetilde{C}) \leqq \varepsilon;$$

i.e. $\nu(O \cdot \widetilde{D}) \leqq \nu(C \cdot \widetilde{D}) + \varepsilon$.

18.5. *The following two conditions on the measure ν are equivalent to each other and to* 18.1 *as well as to* 18.2.

18.5.1. *For all open sets O with compact closure*

$$\nu(O) = \sup \nu(C), \quad C \leqq O,\ C \text{ compact}$$

(*i.e.* 18.1 *holds*).

18.5.2. *For all compact sets C*

$$\nu(C) = \inf \nu(O), \quad O \geqq C, O \text{ open}$$

(*i.e.* 18.2 *holds*).

PROOF. 18.5.1 is equivalent to 18.1 by 18.4.1, and 18.5.2 is equivalent to 18.2 by 18.4.2. So we must only prove that 18.5.1 and 18.5.2 are equivalent.

18.5.1→18.5.2: Let a compact set C be given. Choose any $\varepsilon > 0$. Choose an open set $O \geqq C$ with compact closure. Then $O \cdot \widetilde{C}$ is also an open set with compact closure. We can find, by assumption, a compact set $C' \leqq O \cdot \widetilde{C}$ with $\nu(C') \geqq \nu(O \cdot \widetilde{C}) - \varepsilon$. Thus $\nu(O \cdot \widetilde{C} \cdot \widetilde{C'}) = \nu(O \cdot \widetilde{C}) - \nu(C') \leqq \varepsilon$.

Now $O \cdot \widetilde{C'}$ is open and $O \cdot \widetilde{C'} \geqq O \cdot (\widetilde{O \cdot \widetilde{C}}) = C$. Furthermore

$$\nu(O \cdot \widetilde{C'}) - \nu(C) = \nu(O \cdot \widetilde{C'} \cdot \widetilde{C}) \leqq \varepsilon;$$

i.e. $\nu(O \cdot \widetilde{C'}) \leqq \nu(C) + \varepsilon$.

18.5.2→18.5.1: Let an open set O with compact closure be given. Choose any $\varepsilon > 0$. O's closure, $\overline{O}$, is compact, and with it $\overline{O} \cdot \widetilde{O}$ (the boundary of O). We can find, by assumption, an open set $O' \geq \overline{O} \cdot \widetilde{O}$ with $\nu(O') \leqq \nu(\overline{O} \cdot \widetilde{O}) + \varepsilon$. Thus $\nu(O' \cdot (\widetilde{\overline{O} \cdot \widetilde{O}})) = \nu(O') - \nu(\overline{O} \cdot \widetilde{O}) \leqq \varepsilon$.

Now $\overline{O}$ is compact, and with it $\overline{O} \cdot \widetilde{O'}$. Clearly $\overline{O} \cdot \widetilde{O'} \leqq \overline{O} \cdot (\widetilde{\overline{O} \cdot \widetilde{O}}) = O$. Also, since

$$O \cdot (\widetilde{\overline{O} \cdot \widetilde{O'}}) = O \cdot O' \leqq O' \cdot (\widetilde{\overline{O}} + O) = O' \cdot (\widetilde{\overline{O} \cdot \widetilde{O}}),$$

therefore

$$\begin{aligned} \nu(O) - \nu(\overline{O} \cdot \widetilde{O'}) &= \nu(O \cdot (\widetilde{\overline{O} \cdot \widetilde{O'}})) \\ &\leqq \nu(O' \cdot (\widetilde{\overline{O} \cdot \widetilde{O}})) \leqq \varepsilon; \end{aligned}$$

i.e. $\nu(\overline{O} \cdot \widetilde{O'}) \geqq \nu(O) - \varepsilon$.

18.6. Our results so far motivate the following definition. If ν satisfies any one of the four equivalent conditions of 18.5 we call ν a *regular* measure. In all the following sections we shall assume, in addition to the general assumptions formulated at the beginning of this section, that the measures we are dealing with are regular.

18.7. We remark that 18.5.2 implies 18.5.1 even if we do not require in advance that every compact set has finite measure, but replace this by the weaker assumption that no point has infinite measure. For then 18.5.2 implies that every point p has a neighborhood O_p of finite measure, so that for every compact set C we have

$$C \leqq \sum_{p \in C} O_p.$$

The definition of compactness implies $C \leqq O_{p_1} + \cdots + O_{p_n}$, whence $\nu(C) < \infty$ and the proof may be carried out as in 18.5.

18.8. The considerations of regularity of measures are usually obscured when, in discussing measures, only separable spaces are discussed, since in a locally compact *separable* space every measure is regular. For in such a space every closed set M is easily seen to be the intersection of a countable set of open sets $O_1, O_2, \ldots$ (i.e. every closed set is a G_δ). If $M = C$ is even compact, we can find an open set $O \geqq C$ with compact closure $\overline{O} = D$, and replacing each O_i by $O \cdot O_1 \cdots O_i$, we obtain $D \geqq O_1 \geqq O_2 \geqq \cdots$, and $O_1 O_2 \cdots = C$. Hence $\lim \nu(O_i) = \nu(O_1 O_2 \cdots) = \nu(C)$, and this implies 18.5.2, i.e. the regularity of ν.

The reader who knows Carathéodory's notion of regularity[1] will note that the form 18.2 of our definition of regularity is very similar to Carathéodory's. Nevertheless there is a fundamental difference between the two, which is particularly clear from the above remark. Our regularity is vacuous in a separable space while Carathéodory's original concept was formed in a separable space—n-dimensional Euclidean space. This is, of course, due to our considering Borel sets only, while in this part of Carathéodory's theory the main emphasis is on nonmeasurable sets, which are entirely outside the domain of our discussion.

[1]C. Carathéodory, *Vorlesungen über reelle Funktionen*, Leipzig, 1918; particularly pp. 258 274.

19. Fubini's theorem

In this section we assume that S and T are locally compact topological spaces and that μ and ν are measures defined on the Borel sets of S and T respectively. (We recall that in accordance with the conventions of the preceding paragraph μ and ν are assumed to be regular and to be finite for compact sets.)

In the considerations which follow we shall have to make use of the notion of integration, for both measures μ and ν. But we shall need it only with integrands which are everywhere $\geqq 0$. Accordingly we shall introduce the notion of a *Baire function*, always bearing in mind the two following points: First, that we can restrict the values of our functions to real numbers $\geqq 0$; second, that under our present definitions the entire space may not be a Borel set. (Cf. the beginning of §18 together with 17.4.) Indeed, one concludes immediately from 17.10 that the entire space is a Borel set, in our present sense, if and only if it is the sum of countably many compact sets—which need not be the case. Consequently it is advisable to require that at least the sets where our functions are > 0 (i.e. $\neq 0$) be Borel sets.

For the above reasons we define as follows:

19.1. *A function $f(x)$ is a Baire function if its values are real numbers $\geqq 0$, and if for every $\alpha > 0$ the set of all x for which $f(x) \geqq \alpha$ is a Borel set.*

It is well known that this family of functions is closed under the operations of addition, subtraction, multiplication, and passage to the limit (of an everywhere convergent sequence), and that a theory of integration, analogous in all respects to the usual Lebesgue theory, can be developed in it for any measure of the type we are considering, and we shall freely make use, in what follows, of the concepts of integral, integrability of a Baire function, etc. (Since the functions under consideration are everywhere $\geqq 0$, integrals with the value $+\infty$ are admissible.)

19.2. Let $S \times T$ be the product space of S and T (cf. for example 14.1). For every set $M \leqq S \times T$ we denote by M_x (or M_y) the set of all points $y \in T$ (or $x \in S$) for which $(x, y) \in M$. We denote by $\pi_S(x, y)$ (or $\pi_T(x, y)$) the function $\pi_S(x, y) = x$ (or $\pi_T(x, y) = y$); π_S and π_T are the projections of $S \times T$ on S and T respectively. For any set $M \leqq S \times T$ we write $M_S = \pi_S(M)$, $M_T = \pi_T(M)$. We note that π_S and π_T are continuous functions with domain $S \times T$ and ranges S and T respectively.

19.3. *We establish now the following facts*:

19.3.1. *If M ($\leqq S \times T$) is compact, then M_S, M_T are also compact.*

19.3.2. *M ($\leqq S \times T$) has a compact closure if and only if M_S, M_T have compact closures.*

PROOF. For 19.3.1: $M_S = \pi_S(M)$, $M_T = \pi_T(M)$ are continuous images of the compact set M; hence they are compact by 5.3.

For 19.3.2: Necessity: $\overline{M}$ is compact, so $(\overline{M})_S, (\overline{M})_T$ are compact by 19.3.1. Now $M \leqq \overline{M}$ so $M_S \leqq (\overline{M})_S$, $M_T \leqq (\overline{M})_T$; hence $\overline{M_S}, \overline{M_T}$ are closed subsets of $(\overline{M})_S, (\overline{M})_T$ respectively. Thus they are compact too.

Sufficiency: If $\overline{M_S}, \overline{M_T}$ are compact, then $\overline{M_S} \times \overline{M_T}$ is compact by Tychonoff's theorem. Consequently it is also closed by 5.1. Now clearly $M \leqq M_S \times M_T \leqq \overline{M_S} \times \overline{M_T}$; hence M is a closed subset of $\overline{M_S} \times \overline{M_T}$. Thus it is compact too.

19.4. The principal object of the present section is to prove the following theorem.

Fubini's theorem. For every Borel set $M \leqq S \times T$

19.4.1. M_x *and* M_y *are Borel sets (in* T *and* S*, respectively) for all* x *and* y*;*

19.4.2. $\nu(M_x)$ *and* $\mu(M_y)$ *are Baire functions (defined on* S *and* T*, respectively);*

19.4.3. $\int_S \nu(M_x)\, d\mu(x) = \int_T \mu(M_y)\, d\nu(y)$;

19.4.4. *The common value* $\rho(M)$ *of the integrals in* 19.4.3 *is a completely additive, nonnegative, regular set function defined for all Borel sets* $M \leqq S \times T$ *and finite for all compact sets: in other words* $\rho(M)$ *is a measure in* $S \times T$.

The proof will depend on several auxiliary results which we state separately as lemmas.

19.5. *If* M *is any compact set* $M \leqq S \times T$*, then corresponding to every point* $y_0 \in T$ *and every open set* $O \leqq S$ *for which* $M_{y_0} \leqq O$ *we may find a neighborhood* P *of* y_0*,* $y_0 \in P \leqq T$*, such that* $y \in P$ *implies* $M_y \leqq O$*. In other words* M_y *is an upper semi-continuous function of* y.

PROOF. Let $O' \leqq T$ be an arbitrary neighborhood of y_0, and form the direct product $O \times O'$. For any point $y \in M_T$, $y \neq y_0$, let Q_y be a neighborhood of y, such that $y_0 \notin \overline{Q_y}$. Let O^* be an open set, $O^* \leqq S$ such that $M_S \leqq O^*$. We have

$$M \leqq O \times O' + \sum_{y \in M_T, y \neq y_0} O^* \times Q_y.$$

Since M is compact we may find a finite number of points $y_1, \ldots, y_n \in M_T$ such that

$$M \leqq O \times O' + \sum_{i=1}^{n} O^* \times Q_{y_i}.$$

Let P be the complement of $\sum_{i=1}^n \overline{Q}_{y_i}$; then P is an open set, $P \leqq T$, and, because of the choice of Q_y, $y_0 \in P$. We assert that P is the neighborhood whose existence the theorem states. For if $y \in P$ and $(x, y) \in M$, then $(x, y) \notin O^* \times Q_{y_i}$ for any $i = 1, \ldots, n$ (since $y \notin \overline{Q}_{y_i}$), so that $(x, y) \in O \times O'$, i.e. $x \in O$. Since this is true for all such x, we have $M_y \leqq O$, as was to be proved.

19.6. By a *rectangle* we mean a set of the form $A \times B$ where A and B are Borel sets (in S and T respectively) with finite measures. (Observe that if S and T both coincide with a Euclidean line, hence if $S \times T$ is a Euclidean plane, then our notion of a rectangle is more general than that one of elementary geometry.) By a *rectangular set* we mean a finite or countable sum of pairwise disjoint rectangles. We observe that 19.4.1–19.4.3 are valid for any rectangle (and therefore for any rectangular set). For if $M = A \times B$, then

$$M_x = \begin{cases} B & \text{if } x \in A, \\ \theta & \text{if } x \notin A, \end{cases} \qquad M_y = \begin{cases} A & \text{if } y \in B, \\ \theta & \text{if } y \notin B, \end{cases}$$

so that

$$\nu(M_x) = \begin{cases} \nu(B) & \text{if } x \in A, \\ 0 & \text{if } x \notin A, \end{cases} \qquad \mu(M_y) = \begin{cases} \mu(A) & \text{if } y \in B, \\ 0 & \text{if } y \notin B. \end{cases}$$

Consequently 19.4.1 and 19.4.2 are obvious and

$$\int_S \nu(M_x)\,d\mu(x) = \mu(A)\nu(B) = \int_T \mu(M_y)\,d\nu(y).$$

19.7. *If E is any Borel set in T and if to each $y \in E$ there corresponds a neighborhood O_y of y (so that $E \leqq \sum_{y\in E} O_y$), then there exists a countable sequence $y_1, y_2, \ldots$ of points of E such that $E \leqq \sum_{i=1}^{\infty} Q_{y_i} + N$ where $\nu(N) = 0$.*

PROOF. It is sufficient to prove the theorem in the case when $\nu(E) < \infty$; the general case follows from the fact that every Borel set is the sum of countably many Borel sets of finite measure.

If E has finite measure then, because of regularity, we may find a compact set $C_n \leqq E$ such that $\nu(\widetilde{C}_n E) < 1/n$. Since $C_n \leqq \sum_{y \in C_n} O_y$, there exist a finite number of points $y_1^n, \ldots, y_{k_n}^n$ in C_n (and therefore in M) such that $C_n \leqq \sum_{i=1}^{k_n} O_{y_i^n}$. We write $C = \sum_{n=1}^{\infty} C_n$. Then $\nu(\widetilde{C}E) \leqq \nu(\widetilde{C}_n E) < 1/n$, so that $\nu(\widetilde{C}E) = 0$, and $CE \leqq \sum_{n=1}^{\infty} \sum_{i=1}^{k_n} O_{y_i^n}$; in other words the sets $O_{y_i^n}$ cover E except possibly for a subset of the set $\widetilde{C}E$ of measure zero.

19.8. We are now ready to give the proof of Fubini's theorem. We assume first that M is a compact set, $M \leqq S \times T$.

19.8.1. For any fixed $x \in S$, consider the set of all points $(x, y) \in S \times T$. This set is clearly closed; hence its intersection with M is compact. If we denote this intersection by M_x', then the projection of M_x' on T, $(M_x')_T$ (cf. 19.3.1), is again compact; but $(M_x')_T = M_x$. Hence if M is compact, M_x is certainly a Borel set for every x, and similarly M_y is a Borel set for every y.

19.8.2. If M is compact, $\mu(M_y)$ is an upper semi-continuous function of y for $y \in T$. For any $\varepsilon > 0$ and $y_0 \in T$ we may find an open set $O \leqq S$ such that $M_{y_0} \leqq O$ and $\mu(O) \leqq \mu(M_{y_0}) + \varepsilon$. According to 19.5 we may then find a neighborhood P of y_0, $P \leqq T$, such that for $y \in P$, $M_y \leqq O$ so that $\mu(M_y) \leqq \mu(M_{y_0}) + \varepsilon$. This proves the upper semi-continuity of $\mu(M_y)$.

19.8.3. We can now prove that $\nu(M_x)$ and $\mu(M_y)$ are Baire functions of x and y respectively. By symmetry we may restrict ourselves to considering $\mu(M_y)$.

Consider an $\alpha > 0$. Since $\mu(M_y)$ is upper semi-continuous by 19.8.2, the set of all y with $\mu(M < y) < \alpha$ is open. Hence the complementary set of all y with $\mu(M_y) \geqq \alpha$ is closed. This set is a subset of the compact set M_T, since $y \notin M_T$ implies $M_y = \theta$, $\mu(M_y) = 0$, so that the set of all y with $\mu(M_y) \geqq \alpha$ is compact, and therefore a Borel set. Thus $\mu(M_y)$ is a Baire function of y.

19.8.4. We prove next (assuming still that M is compact) that for any $\delta > 0$ we may find a rectangular set K (cf. 19.6) such that $M \leqq K$ and

$$\int_T \mu(K_y)\,d\nu(y) \leqq \int_T \mu(M_y)\,d\nu(y) + \delta.$$

Let ε be an arbitrary positive number. The compactness of M implies that of M_S, so that $\mu(M_S) < \infty$; we may find a positive integer k such that $\mu(M_S) < k\varepsilon$. For every $i = 0, 1, \ldots, k-1, k$, we write M_T^i for the set of points $y \in M_T$ for which $\mu(M_y) \geqq i\varepsilon$. Then we have

$$M_T = M_T^0 \geqq M_T^1 \geqq \cdots \geqq M_T^{k-1} \geqq M_T^k = \theta \tag{19.8.4.1}$$

and every M_T^i is a Borel set (in fact, according to 19.8.3, M_T^i is compact).

For each $y \in M_T$ let O_y be an open set ($O_y \leqq S$) such that $M_y \leqq O_y$ and $\mu(O_y) \leqq \mu(M_y) + \varepsilon$. Then (cf. 19.5) we may find a neighborhood P_y of y such that $y' \in P_y$ implies $M_{y'} \leqq O_y$. It follows that for $i = 1, \ldots, k$,

$$M_T^{i-1} \cdot \widetilde{M_T^i} \leqq \sum P_y, \qquad y \in M_T^{i-1} \cdot \widetilde{M_T^i},$$

whence, by 19.7, we may find a sequence of points $y_n^i \in M_T^{i-1} \cdot \widetilde{M_T^i}$ and a set N_i with $\nu(N_i) = 0$ such that

$$M_T^{i-1}\widetilde{M_T^i} \leqq \sum_{n=1}^{\infty} P_{y_n^i} + N_i. \tag{19.8.4.2}$$

We observe that the sets $P_{y_n^i}$ and N_i have three definitory properties:

(i) $P_{y_n^i}$ is an open set.

(ii) $y' \in P_{y_n^i}$ implies $M_{y'} \leqq O_{y_n^i}$.

(iii) $\nu(N_i) = 0$.

We have used (i) in order to be able to apply 19.7, to derive (19.8.4.2). From now on we shall only use (19.8.4.2), while (i) itself will not be needed any more. (ii), (iii) however will be necessary.

We shall replace the sets $P_{y_n^i}$ and N_i by certain subsets. Thus (ii), (iii) will automatically remain true, while (i) will be lost. The fate of (19.8.4.2) must be watched—as a matter of fact we shall replace it by stronger statements.

Replace first every $P_{y_n^i}$ by its subset $P_{y_n^i} \cdot (\widetilde{\sum_{m=1}^{n-1} P_{y_m^i}})$, and N_i by its subset $N_i \cdot (\widetilde{\sum_{m=1}^{\infty} P_{y_m^i}})$. Thus (i) is lost, while (ii), (iii) remain true. $\sum_{n=1}^{\infty} P_{y_n^i} + N_i$ is not changed; hence (19.8.4.2) remains true. And the $P_{y_n^i}$ ($n = 1, 2, \ldots$) and N_i are now pairwise disjoint. (We consider one fixed i.)

Next replace every $P_{y_n^i}$ by its subset $M_T^{i-1} \cdot \widetilde{M_T^i} \cdot P_{y_n^i}$ and N_i by its subset $M_T^{i-1} \cdot \widetilde{M_T^i} \cdot N_i$. Thus (ii), (iii) are still true, the $P_{y_n^i}$ ($n = 1, 2, \ldots$) and N_i are still disjoint (for one fixed i), and in (19.8.4.2) the $\leqq$ is replaced by $=$. We restate the latter:

$$M_T^{i-1}\widetilde{M_T^i} = \sum_{n=i}^{\infty} P_{y_n^i} + N_i. \tag{19.8.4.3}$$

Summing on i ($= 1, \ldots, k$) we obtain, remembering (19.8.4.1),

$$M_T = \sum_{i=1}^{k} \sum_{n=1}^{\infty} P_{y_n^i} + N_i, \tag{19.8.4.4}$$

where $N = \sum_{i=1}^{k} N_i$.

(19.8.4.3) and (19.8.4.1) show that the $P_{y_n^i}$, N_i of different i are also disjoint, so we see:

(iv) All $P_{y_n^i}$ ($i = 1, \ldots, k, n = 1, 2, \ldots$) and N are pairwise disjoint. And (iii) gives

(v) $\nu(N) = 0$.

So we have at present (19.8.4.3), (19.8.4.4) and (ii), (iii), (iv), (v).

Consider now the set (in $S \times T$)

$$K = \sum_{i=1}^{k} \sum_{n=1}^{\infty} O_{y_n^i} \times P_{y_n^i} + M_S \times N. \tag{19.8.4.5}$$

If $(x,y) \in M$, then $y \in M_{T'}$, so that by (19.8.4.4) either $y \in P_{y_n^i}$ for some $i = 1,\dots,k$, $n = 1,2,\dots$, or $y \in N$. If $y \in P_{y_n^i}$, then (ii) gives $x \in rO_{y_n^i}$; hence $(x,y) \in O_{y_n^i} \times P_{y_n^i}$. If $y \in N$, then observe that at any rate $x \in M_S$; hence $(x,y) \in M_S \times N$. Consequently we have in all cases $(x,y) \in K$, i.e.

$$M \leqq K. \tag{19.8.4.6}$$

By (19.8.4.5) and (iv), K is a rectangular set, and (19.8.4.6) establishes $M \leq K$. So we must only evaluate $\int_{T'} \mu(K_y)\, d\nu(y)$.

Now we have, considering (19.8.4.5) and (iv),

$$K_y = \begin{cases} O_{y_n^i} & \text{for } y \in P_{y_n^i}, \\ M_S & \text{for } y \in N, \\ \theta & \text{otherwise.} \end{cases}$$

Hence (iv), (v) give

$$\begin{aligned} \int_{T'} \mu(K_y)\, d\nu(y) &= \sum_{i=1}^{k} \sum_{n=1}^{\infty} \mu(O_{y_n^i}) \nu(P_{y_n^i}) \\ &\leqq \sum_{i=1}^{k} \sum_{n=1}^{\infty} (\mu(M_{y_n^i}) + \varepsilon) \nu(P_{y_n^i}), \end{aligned}$$

since $y_n^i \in M_{T'}^{i-1} \cdot \widetilde{M_{T'}^i}$, and since $y \in M_{T'}^i$ implies $\mu(M_y) < i\varepsilon$ we obtain finally

$$\int_{T'} \mu(K_y)\, d\nu(y) \leqq \sum_{i=1}^{k} \sum_{n=1}^{\infty} (i+1)\varepsilon\nu(P_{y_n^i}). \tag{19.8.4.7}$$

On the other hand, by (19.8.4.4) and (iv), (v),

$$\int_{T'} \mu(M_y)\, d\nu(y) = \sum_{i=1}^{k} \sum_{n=1}^{\infty} \int_{P_{y_n^i}} \mu(M_y)\, d\nu(y).$$

Since $P_{y_n^i} \leqq M_{T'}^{i-1} \cdot \widetilde{M_{T'}^i}$ by (19.8.4.3), and since $y \in M_{T'}^{i-1}$ implies $\mu(M_y) \geqq (i-1)\varepsilon$ we obtain here

$$\int_{T'} \mu(M_y)\, d\nu(y) \geqq \sum_{i=1}^{k} \sum_{n=1}^{\infty} (i-1)\varepsilon\nu(P_{y_n^i}). \tag{19.8.4.8}$$

Subtracting (19.8.4.8) from (19.8.4.7), and using (19.8.4.4) and (iv), (v), we obtain

$$\begin{aligned} &\int_{T'} \mu(K_y)\, d\nu(y) - \int_{T'} \mu(M_y)\, d\nu(y) \\ &\qquad \leqq 2\varepsilon \sum_{i=1}^{k} \sum_{n=1}^{\infty} \nu(P_{y_n^i}) = 2\varepsilon\nu(M_{T'}). \end{aligned}$$

Hence choosing $\varepsilon = \delta / 2\nu(M_{T'})$ gives

$$\int_{T'} \mu(K_y)\, d\nu(y) - \int_{T'} \mu(M_y)\, d\nu(y) \leqq \delta, \tag{19.8.4.9}$$

which is the desired inequality.

19.8.5. Let δ be an arbitrary positive number, and M a compact set, $M \leqq S \times T$. Then 19.8.4 enables us to find a rectangular set $K \geqq M$, such that $\int_T \mu(K_y)\,d\nu(y) \leqq \int_T \mu(M_y)\,d\nu(y) + \delta$. Using the property of rectangular sets established in 19.6 we obtain

$$\int_N (M_x)\,d\mu(x) \leqq \int \nu(K_x)\,d\mu(x)$$
$$= \int \mu(K_y)\,d\nu(y) \leqq \int \mu(M_y)\,d\nu(y) + \delta,$$

whence, since this is true for all $\delta > 0$

$$\int \nu(M_x)\,d\mu(x) \leqq \int \mu(M_y)\,d\nu(y).$$

The symmetry of the roles of S and T implies the opposite inequality, thus finally proving the validity of 19.4.3 for compact sets M.

19.8.6. In 19.8.3 and 19.8.4 we proved that 19.4.1–19.4.3 are valid for all compact sets M. Their validity for arbitrary Borel sets follows from 17.11 since it is immediately verified that the properties 19.4.1–19.4.3 are hereditary under all steps of the construction there given. Cf. the detailed discussion of the constructions involved in these steps, as given after (17.11). Thus only the assertion 19.4.4 concerning the properties of the resulting set function $\rho(M)$ remains. That $\rho(M)$ is nonnegative and completely additive is clear; the only thing that remains to be proved is that $\rho(M)$ is regular. This in turn is an easy consequence of 19.8.4 and 18.5.2. According to 18.5.2 it is sufficient to prove that every compact set can be approximated arbitrarily closely from above by open sets. According to 19.8.4 every compact set can be approximated by rectangular sets, so that it is only necessary to prove the approximability of rectangular sets by open sets. This, however, is an immediate consequence of the same assertion for rectangles, considering 19.6. For rectangles it follows at once from the regularity of the given measures, μ and ν and from the definition of topology in the product space $S \times T$.

This concludes the proof of Fubini's theorem.

20. Uniqueness of Haar measure

We return to the consideration of groups. As usual we assume that G is a locally compact topological group and that $\nu(M)$ is a left-invariant regular measure defined on the Borel sets of G, which is finite for compact sets and positive for open sets. The principal object of this section is to prove that, up to a multiplicative constant factor, ν is uniquely determined by these requirements; in other words *if μ is any left-invariant regular measure defined on all Borel sets of G and finite for compact sets, then there exists a positive, finite constant c such that $\mu(M) = c\nu(M)$ for all Borel sets M.* The proof of this statement will depend on several auxiliary theorems which are of interest in themselves.

20.1. Throughout what follows we shall make use of the direct product group $G \times G$ and in it the measure $\rho(M)$ obtained by applying Fubini's theorem:

(20.1.1) $$\rho(M) = \int \nu(M_x)\,d\nu(x) = \int \nu(M_y)\,d\nu(y).$$

We assert, first, that the following one-to-one mappings of $G \times G$ into itself are all measure preserving:

(20.1.2) $$(x, y) \to (ax, y)$$

(20.1.3) $$(x, y) \to (x, by),$$

(20.1.4) $$(x, y) \to (ax, by)$$

(20.1.5) $$(x, y) \to (y, x),$$

(20.1.6) $$(x, y) \to (x, xy),$$

(20.1.7) $$(x, y) \to (x, x^{-1}y).$$

For (20.1.2) we see this from the relation $\rho(M) = \int \nu(M_y)\, d\nu(y)$. If we denote by M' the image of M under the transformation $(x, y) \to (ax, y)$, then M'_y is the set of all x for which $(x, y) \in M'$, or, equivalently, the set of all ax for which $(ax, y) \in M'$. The latter set, on the other hand, is an M_y. Hence

$$\rho(M') = \int \nu(M'_y)\, d\nu(y) = \int \nu(aM_y)\, d\nu(y) = \int \nu(M_y)\, d\nu(y) = \rho(M),$$

as was to be proved.

For (20.1.3) the result follows just as above, using, this time, the relation $\rho(M) = \int \nu(M_x)\, d\nu(x)$. The product of the two mappings (20.1.2) and (20.1.3) is (20.1.4), so that the latter is also measure preserving. We observe that the assertion that $\rho(M)$ is invariant under (20.1.4) is merely the assertion that $\rho(M)$ is a left invariant Haar measure in $G \times G$. The measure preserving character of (20.1.5) is immediate from the symmetry, in x and y, of (20.1.1).

Finally, for (20.1.6), we prove, as above, that $M'_x = xM_x$ (where M' is the image of M under (20.1.6)), and (20.1.7) is the inverse mapping of (20.1.6).

Applying (20.1.6), (20.1.5), and (20.1.7), in this order, we obtain the fact that

(20.1.8) $$(x, y) \to (xy, y^{-1})$$

is a measure-preserving mapping of $G \times G$ into itself. Hence if M is any Borel set in G, and we apply this mapping to the set $G \times M$, we obtain $G \times M^{-1}$. Therefore $\rho(G \times M) = \rho(G \times M^{-1})$, i.e.

(20.1.9) $$\nu(G)\nu(M) = \nu(G)\nu(M^{-1}).$$

Observe that if such a product $\rho\ (A \times B) = \nu(A)\nu(B)$ is of the form $0 \cdot \infty$ or $\infty \cdot 0$ its value must be taken to be 0. This is clear from its origin in the integral formulae of 19.6. Observe furthermore that certainly $v(G) > 0$, but $v(G) \leqq \infty$.

20.2. From the above results we conclude for every fixed choice of $\nu(M)$ (subject to the requirements for $\nu(M)$ at the beginning of §20):

20.2.1. *If G is compact, then $\nu(M)$ is both right and left invariant and also inverse invariant; i.e.*

(20.2.1.1) $$\nu(aM) = \nu(M),$$

(20.2.1.2) $$\nu(Ma) = \nu(M),$$

(20.2.1.3) $$\nu(M^{-1}) = \nu(M).$$

20.2.2. *For every G the vanishing of $\nu(M)$ is both right and left invariant and also inverse invariant; i.e.*

$$\nu(aM) = 0 \quad \text{is equivalent to} \quad \nu(M) = 0, \tag{20.2.2.1}$$

$$\nu(Ma) = 0 \quad \text{is equivalent to} \quad \nu(M) = 0, \tag{20.2.2.2}$$

$$\nu(M^{-1}) = 0 \quad \text{is equivalent to} \quad \nu(M) = 0. \tag{20.2.2.3}$$

PROOF. Ad 20.2.1: By 16.2, $\nu(G) < \infty$; hence (20.1.9) gives (20.2.1.3). (20.2.1.1) holds by definition. (20.2.1.2) ensues by application, in this order, of (20.2.1.3), (20.2.1.1) (with a^{-1}), and (20.2.1.3).

Ad 20.2.2: (20.1.9) still permits us to infer (20.2.2.3). (Cf. the remark following (20.1.9).) (20.2.2.1) holds by definition. (Even (20.2.1.1) holds.) (20.2.2.2) ensues by application, in this order, of (20.2.2.3), (20.2.2.1) (with a^{-1}), and (20.2.2.3).

20.3. We establish now a relation to which we shall refer as the *average theorem*:

If M and N are Borel sets in G, then

$$\int_G \nu(M \cdot xN)\, d\nu(x) = \nu(M)\nu(N^{-1}).$$

PROOF. Form the set $E = M \times N^{-1} \leqq G \times G$; then the image E' of E under the (measure-preserving) mapping (20.1.6) $(x,y) \to (x,xy)$ is described as follows: $(x',y') \in E'$ means that $x' = x$ is arbitrary in M and $y' = xy$ for $y \in N^{-1}$: in other words E' is the set of all (x,y) with $x \in M$, $y \in xN^{-1}$. Hence for all y, E'_y is the set of all x for which $x \in M$ and $x \in yN$:

$$E'_y = M \cdot xN.$$

From the assertion that the mapping $(x,y) \to (x,xy)$ is measure preserving, we obtain immediately the desired relation:

$$\rho(E) = \nu(M)\nu(N^{-1}) = \rho(E') = \int \nu(E'_y)\, d\nu(y) = \int \nu(M \times N)\, d\nu(y).$$

20.4. As a consequence of the average theorem we obtain the following result:

If for any Borel set M, $\nu(\widetilde{M} \cdot xM) = 0$ for all x, then either $\nu(M) = 0$ or $\nu(\widetilde{M}) = 0$.

PROOF. Assume $\nu(\widetilde{M} \cdot xM) = 0$ for all x. Then by virtue of the average theorem

$$0 = \int_G \nu(\widetilde{M} \cdot xM)\, d\nu(x) = \nu(\widetilde{M})\nu(M^{-1});$$

hence either $\nu(\widetilde{M}) = 0$ or $\nu(M^{-1}) = 0$, and therefore, by (20.2.2.3), $\nu(M) = 0$.

This theorem asserts that a locally compact topological group, considered as a group of measure-preserving transformations on itself, is ergodic.

20.5. *If M and N are any two Borel sets with compact closures and positive measures, then there exist a Borel set K and elements x, y such that $xK \leqq M$, $yK \leqq N$, and*

$$\nu(K) \geqq \frac{\nu(M)\nu(N^{-1})}{\nu(M \odot N^{-1})},$$

$\nu(M), \nu(N), \nu(M \odot N^{-1})$ being all positive and finite.

PROOF. $\nu(M), \nu(N)$ we assumed to be positive—they are finite since $\overline{M}$ and $\overline{N}$ are compact. Thus $M \neq \theta$ and we may choose an $a \in M$. Since $\nu(N)$ is positive, so is $\nu(N^{-1})$; hence

$$\nu(M \odot N^{-1}) \geqq \nu(aN^{-1}) = \nu(N^{-1}) > 0.$$

Moreover $M \odot N^{-1} \leqq \overline{M} \odot \overline{N}^{-1}$, which is compact. Hence $\nu(M \odot N^{-1})$ is finite.

$$x \notin M \odot N^{-1} \quad \text{implies} \quad M \cdot xN = \theta, \nu(M \cdot xN) = 0,$$

so that, applying the average theorem,

$$\int_{M \odot N^{-1}} \nu(M \times N)\, d\nu(x) = \int_G \nu(M \times N)\, d\nu(x) = \nu(M)\nu(N^{-1}).$$

Hence for at least one $x_0 \in M \odot N^{-1}$ (in fact for x_0 in some subset of positive measure in $M \odot N^{-1}$) we must have

$$\nu(M \cdot x_0 N) \geqq \frac{\nu(M)\nu(N^{-1})}{\nu(M \odot N^{-1})}.$$

Choosing $K = M \cdot x_0 N$, $x = 1$, $y = x_0^{-1}$, establishes the desired result.

Using 20.5 we shall now establish the principal result needed for the proof of the uniqueness of Haar measure: a decomposition theorem valid for any left-invariant Haar measure ν.

20.6. *If M and N are any two Borel sets in G, there exist in G two sequences of elements x_n, y_n, $n = 1, 2, \ldots$, and a sequence of Borel sets $K_1, K_2, \ldots$ such that*

(20.6.1) $$M = x_1 K_1 + x_2 K_2 + \cdots + M_\infty,$$

(20.6.2) $$N = y_1 K_1 + y_2 K_2 + \cdots + N_\infty,$$

where the addends in each sum are pairwise disjoint, and either $\nu(M_\infty) = 0$ or $\nu(N_\infty) = 0$.

PROOF. Since every Borel set can be written as the disjoint sum of a countable sequence of Borel sets with compact closures, it is sufficient to prove the theorem for the case in which both M and N have compact closures.

We write $M_0 = M$, $N_0 = N$. If M_i, N_i have been defined for $i = 0, \ldots, n$, so that $\overline{M}_i$ and $\overline{N}_i$ are compact, and $\nu(M_i) > 0$, $\nu(N_i) > 0$, $i = 0, \ldots, n$, we apply 20.5 and obtain x_n, y_n, K_n so that $x_n K_n \leqq M_n$, $y_n K_n \leqq N_n$, and

$$\nu(K_n) \geqq \frac{\nu(M_n)\nu(N_n^{-1})}{\nu(M_n \odot N_n^{-1})}.$$

Then we define

$$M_{n+1} = M_n \cdot \widetilde{x_n K_n}, \qquad N_{n+1} = N_n \cdot \widetilde{y_n K_n}.$$

If for any positive integer n, $\nu(M_n) = 0$, we write $K_n = K_{n+1} = \cdots = \theta$, $M_\infty = M_n$, $N_\infty = N_n$, and, with arbitrary choice of $x_n, x_{n+1}, \ldots, y_n, y_{n+1}, \ldots$ we have the desired decomposition. Hence we may suppose that the induction continues indefinitely. In this case we write $M_\infty = \prod_{n=0}^\infty M_n$, $N_\infty = \prod_{n=0}^\infty N_n$, and we obtain decompositions of the form (20.6.1) and (20.6.2). It remains to prove only that at least one of the sets M_∞ and N_∞ has measure zero.

Since

$$M = M_0 \geqq M_1 \geqq \cdots, \qquad N = N_0 \geqq N_1 \geqq \cdots,$$

$\nu(M_n)$ and $\nu(N_n)$ are monotone decreasing sequences; we write $\alpha = \lim_{n\to\infty} \nu(M_n)$, $\beta = \lim_{n\to\infty} \nu(N_n^{-1})$. Then

$$\begin{aligned}\nu(M_{n+1}) &= \nu(M_n) - \nu(x_n K_n) = \nu(M_n) - \nu(K_n)\\ &\leqq \nu(M_n) - \frac{\nu(M_n)\nu(N_n^{-1})}{\nu(M_n \odot N_n^{-1})}\\ &\leqq \nu(M_n) - \frac{\nu(M_n)\nu(N_n^{-1})}{\nu(M \odot N^{-1})},\end{aligned}$$

so that, going to the limit,

$$\alpha \leqq \alpha - \frac{\alpha\beta}{\nu(M \odot N^{-1})}.$$

This implies (since $\alpha \geqq 0$, $\beta \geqq 0$) that at least one of the two numbers α and β is zero; since $\alpha = \nu(M_\infty)$, $\beta = \nu(N_\infty^{-1})$, and since $\nu(N_\infty^{-1})$ vanishes if and only if $\nu(N_\infty)$ does, the desired conclusion follows.

20.7. If M and N are Borel sets with compact closures, then, in the notation of 20.6, $\nu(M) - \nu(N) = \nu(M_\infty) - \nu(N_\infty)$. Hence, since either $\nu(M_\infty) = 0$ or $\nu(N_\infty) = 0$, $\nu(M) >, =, < \nu(N)$ according as $\nu(M_\infty) \neq 0$, $\nu(N_\infty) = 0$, or $\nu(M_\infty) = \nu(N_\infty) = 0$, or $\nu(M_\infty) = 0$, $\nu(N_\infty) \neq \theta$. Since this is valid for all left-invariant Haar measures, we see already that, even without knowing the uniqueness theorem, knowing which sets have measure zero enables us to compare the sizes of any two sets. In other words, if μ and ν are two left-invariant Haar measures such that $\nu(M) = 0$ is equivalent to $\mu(M) = 0$, then the sign of $\nu(M) - \nu(N)$ is the same as the sign of $\mu(M) - \mu(N)$. Another way of stating this result is the following:

Let μ and ν be two left-invariant measures such that $\nu(M) = 0$ is equivalent to $\mu(M) = 0$. Let Δ_μ and Δ_ν denote the ranges of $\mu(M)$ and $\nu(M)$ respectively, as M varies over all Borel sets with compact closure. Then there exists a one-to-one monotone mapping, $\alpha = f(\beta)$ of Δ_ν on Δ_μ such that for every M with compact closure $\mu(M) = f(\nu(M))$.

20.8. *If μ and ν are left-invariant measures such that $\mu(M) = 0$ is equivalent to $\nu(M) = 0$, then there exists a positive finite constant c such that $\mu(M) = c\nu(M)$.*

PROOF. As usual there is no loss of generality in restricting our attention to sets M with compact closure. We use the notations and results of 20.7.

If $\beta_1, \beta_2 \in \Delta_\nu$, $\beta_1 \leqq \beta_2$, then we may find Borel sets M and N with compact closures such that $\nu(M) = \beta_1$, $\nu(N) = \beta_2$. We write M and N in the form (20.6.1) and (20.6.2) respectively, and write $M' = y_1 K_1 + y_2 K_2 + \cdots$. Then $\nu(M') = \nu(M) = \beta_1$ and $M' \leqq N$, so that $N = M' + N'$ with $M' \cdot N' = \theta$. Since $\nu(N) = \beta_2$ and $\nu(M') = \beta_1$, we must have $\nu(N') = \beta_2 - \beta_1$, thus proving that $\beta_2 - \beta_1 \in \Delta_\nu$; moreover since $\mu(N) = f(\beta_2)$, $\mu(M') = f(\beta_1)$, and $\mu(N') = f(\beta_2 - \beta_1)$, we have $f(\beta_2) = f(\beta_1) + f(\beta_2 - \beta_1)$. We restate this:

$$(*) \qquad \begin{cases}\beta_1, \beta_2 \in \Delta_\nu \text{ and } \beta_1 \leqq \beta_2 \text{ imply } \beta_2 - \beta_1 \in \Delta_\nu,\\ \text{and } f(\beta_2 - \beta_1) = f(\beta_2) - f(\beta_1).\end{cases}$$

Now $(*)$ and the nonnegativity of $f(\beta)$ imply

$$(**) \qquad f(\beta) = c\beta \text{ for a positive, finite constant } c.$$

And this is what we wanted to prove. Several ways to establish the implication of (**) by (*) (of which we made use) are known. For the reader who may not be familiar with them, we give a direct and elementary proof of this implication.

20.9. Δ_ν contains positive elements. Assume first that there exists a smallest one among these. Denote it by α_0. For every $\alpha \in \Delta$ there exists an $i = 0, 1, 2, \dots,$ such that $i\alpha_0 \leqq \alpha < (i+1)\alpha_0$. i-fold application of $(*)$ above shows that $\alpha - i\alpha_0 \in \Delta$. Hence $0 < \alpha - i\alpha_0 < \alpha_0$ is impossible. Now clearly $0 \leqq \alpha - i\alpha_0 < \alpha_0$; consequently $\alpha - i\alpha_0 = 0$. Again i-fold application of $(*)$ gives $f(\alpha) - if(\alpha_0) = 0$. Thus $f(\alpha) = if(\alpha_0) = \frac{\alpha}{\alpha_0} f(\alpha_0)$; i.e. $f(a) = c\alpha$ with $c = f(\alpha_0)/\alpha_0$.

Assume next that no smallest positive element of Δ_ν exists. Let α_1 be the greatest lower bound of Δ_ν's positive elements. Clearly $\alpha_1 \geqq 0$. Assume now $\alpha_1 > 0$. Then $\alpha_1 \notin \Delta_\nu$ by our previous assumption. Now $2\alpha_1 > \alpha_1$; so choose $\alpha \notin \Delta_\nu$ with $\alpha < 2\alpha_1$. Owing to the above $\alpha > \alpha_1$; so choose $\beta \in \Delta_\nu$ with $\beta < \alpha$. Again $\beta > \alpha_1$. So $\alpha_1 < \beta < \alpha < 2\alpha_1$. Hence $\alpha - \beta \in \Delta_\nu$ by $(*)$, and $0 < \alpha - \beta < \alpha_1$, contradicting the definition of α_1. Consequently $\alpha_1 = 0$.

Thus for every $\varepsilon > 0$ an $\alpha \in \Delta_\nu$ with $0 < \alpha < \varepsilon$ exists.

Now consider two arbitrary $\beta_1, \beta_2 \in \Delta_\nu$, both positive. We claim that

$$(\#) \qquad \frac{f(\beta_1)}{\beta_1} = \frac{f(\beta_2)}{\beta_2}.$$

Otherwise we may assume, by interchanging β_1, β_2 if necessary, that

$$\frac{f(\beta_1)}{\beta_1} > \frac{f(\beta_2)}{\beta_2}.$$

Consequently

$$(\S) \qquad \frac{f(\beta_1)}{\beta_1} > \frac{(p_0+1)^2}{p_0^2} \cdot \frac{f(\beta_2)}{\beta_2}, \quad \text{for a suitable } p_0 = 1, 2, \dots .$$

Choose an $\alpha \in \Delta_\nu$ with $0 < \alpha < \mathrm{Min}(\frac{\beta_1}{p_0}, \frac{\beta_2}{p_0})$. Choose $r, s = 0, 1, 2, \dots$ with $r\alpha \leqq \beta_1 < (r+1)\alpha, s\alpha \leqq \beta_2 \leqq (s+1)\alpha$. Then $r, s \geqq p_0$. Now repeated application of $(*)$ above gives:

$$f(\beta_2) - sf(\alpha) = f(\beta_2 - s\alpha) \geqq 0,$$
$$(r+1)f(\alpha) - f(\beta_1) = f((r+1)\alpha) - f(\beta_1) = f((r+1)\alpha - \beta_1) \geqq 0;$$

hence

$$f(\beta_1) \leqq (r+1)f(\alpha), \qquad f(\beta_2) \geqq sf(\alpha)$$

and so

$$\frac{f(\beta_1)}{\beta_1} \leqq \frac{(r+1)f(\alpha)}{r \cdot \alpha} = \frac{r+1}{r} \cdot \frac{f(\alpha)}{\alpha} \leqq \frac{p_0+1}{p_0} \frac{f(\alpha)}{\alpha},$$
$$\frac{f(\beta_2)}{\beta_2} \leqq \frac{sf(\alpha)}{(s+1)\alpha} = \frac{s}{s+1} \frac{f(\alpha)}{\alpha} \geqq \frac{p_0}{p_0+1} \frac{f(\alpha)}{\alpha}$$

giving

$$\frac{f(\beta_1)}{\beta_1} \leqq \frac{(p_0+1)^2}{p_0^2} \frac{f(\beta_2)}{\beta_2}.$$

This contradicts (§). Hence (#) is established.

Now let c be the common value of $f(\beta)/\beta$ for all positive $\beta \in \Delta_\nu$ (use (#)). Then $f(\beta) = c\beta$ for all positive $\beta \in \Delta_\nu$, and it is obviously true for $\beta = 0$ (put $\beta_1 = \beta_2$ in $(*)$).

We are now in the position to prove the uniqueness of Haar measure.

20.10. *If μ and ν are left-invariant measures, then there exists a positive constant c such that $\mu(M) = c\nu(M)$ for all Borel sets M.*

PROOF. Again there is no loss of generality in restricting our attention to sets M with compact closure. We consider the left-invariant measures

$$\rho(M) = \nu(M) + \mu(M),$$
$$\sigma(M) = 2\nu(M) + \mu(M).$$

Then $\rho(M) = 0$ and $\sigma(M) = 0$ are equivalent, since they are each equivalent to the simultaneous validity of $\nu(M) = 0$ and $\mu(M) = 0$. Hence, by 20.8, there exists a positive finite constant γ such that

$$\rho(M) = \delta\sigma(M),$$

or

$$(1-\gamma)\mu(M) = (2\gamma - 1)\nu(M).$$

If $\gamma \leqq \frac{1}{2}$, then $\mu(M) \leqq 0$ for all M, so that $\mu(M) \equiv 0$; if $\gamma \geqq 1$, then $\nu(M) \leqq 0$ for all M so that $\nu(M) \equiv 0$. Since we have assumed that μ and ν are positive for open sets, neither of these possibilities can arise. Consequently $\frac{1}{2} < \gamma < 1$, and 20.4 is true with $c = \frac{1-\gamma}{2\gamma-1}$.

21. Consequences

We conclude this chapter by deriving, as a consequence of the uniqueness theorem, certain connections between right and left invariance.

21.1. If $\nu(M)$ is a left-invariant measure, then $\nu(Ma)$ is also a left-invariant measure for every $a \in G$. Hence, by 20.7, we may find a positive finite constant $c = c(a)$ such that

$$\nu(Ma) \equiv c(a)\nu(M).$$

We shall investigate properties of the function $c(a)$.

21.2. $c(ab) = c(a)c(b)$.

21.3. *If a is in the center of G, $c(a) = 1$.*

21.4. *If a is in the commutator subgroup of G, $c(a) = 1$.*

21.5. *$c(a)$ is a continuous function of a.*

PROOF. For 21.2: $\nu(Mab) = c(b)\nu(Ma) = c(b)c(a)\nu(M) = c(ab)\nu(M)$.

For 21.3: If $ab = ba$ for every $b \in G$, then $\nu(Ma) = \nu(aM) = \nu(M)$.

For 21.4: If $a = bdb^{-1}d^{-1}$, then $c(a) = c(b)c(d)\frac{1}{c(b)}\frac{1}{c(d)} = 1$; it follows that $c(a) = 1$ whenever a is a product of commutators.

For 21.5: We note first that if a compact set M is contained in an open set O, $M \leqq O$, we can find a neighborhood P of the identity, so that $M \odot P \leqq O$. For, if $x \in M$, we can, using only the continuity of multiplication, find two neighborhoods Q_x and P_x, $x \in Q_x$, $1 \in P_x$, such that $x' \in Q_x$ and $a' \in P_x$ implies $x'a' \in O$.

Since the family of open sets Q_x, $x \in M$, covers the compact set M, we may find a finite number among them which cover M, $M \leqq Q_{x_1} + \cdots + Q_{x_n}$. Then we may write $P = P_{x_1} \cdots P_{x_n}$; P is the neighborhood of the identity whose existence was asserted. For suppose $x \in M$, $a \in P$. Then $x \in Q_{x_i}$ for some i; since $a \in P_{x_i}$ we have $x \in O$, and since this is true for all $x \in M$ and $a \in P$, we have $M \odot P \leqq O$.

To prove 21.5, let M be a compact set of positive measure, $0 < \nu(M) < \infty$. If ε is any positive number, we may find an open set $O \geqq M$, such that $\nu(O) \leqq \nu(M)(1+\varepsilon)$. Then we may find, according to the preceding paragraph, a neighborhood P of the identity so that $M \odot P \leqq O$. We write $P' = P \cdot P^{-1}$: then P is also a neighborhood of the identity. If $a \in P'$, we have

$$c(a)\nu(M) = \nu(Ma) \leqq \nu(O) \leqq \nu(M)(1+\varepsilon)$$

and

$$\frac{\nu(M)}{c(a)} = \nu(Ma^1) \leqq \nu(O) \leqq \nu(M)(1+\varepsilon),$$

so that $\frac{1}{1+\varepsilon} \leqq c(a) \leqq 1+\varepsilon$. This proves the continuity of $c(a)$ at $a = 1$; the continuity at every point follows as usual from the functional equation 21.2 and the topological homogeneity of the group G.

We remark that the properties 21.2 and 21.5 of $c(a)$ yield still another proof of the identity of the left- and right-invariant measures in case of a compact group. For if G is compact the subset $c(G)$ of the real line is a compact multiplicative subgroup of the positive real numbers and therefore consists of only one point, $c(G) = (1)$. In other words, $c(a) = 1$, so that $\nu(M) = \nu(M)$.

21.6. If ν is a left-invariant measure, let us consider the set functions

$$\mu_1(M) = \nu(M^{-1}) \quad \text{and} \quad \mu_2(M) = \int_M \frac{1}{c(x)}\, d\nu(x).$$

It is easily verified that $\mu_1(M)$ and $\mu_2(M)$ are both *right*-invariant Haar measures. (For $\mu_2(M)$ 21.1 and 21.2 must be applied.) Hence, using the uniqueness theorem for right-invariant measures, we can find a positive finite constant c such that

$$\int_M \frac{1}{c(x)}\, d\nu(x) = c\nu(M^{-1}), \tag{21.6.1}$$

for all Borel sets M. We shall show that $c = 1$. For this purpose we observe that for every Baire function $f(x)$ (21.6.1) implies

$$\int f(x)\frac{1}{c(x)}\, d\nu(x) = c\int f(x^{-1})\, d\nu(x). \tag{21.6.2}$$

Replacing $f(x)$ by $f(x^{-1})$ we obtain

$$\int f(x^{-1})\frac{1}{c(x)}\, d\nu(x) = c\int f(x)\, d\nu(x). \tag{21.6.3}$$

Writing $g(x^{-1}) = f(x^{-1})\frac{1}{c(x)}$ and applying (21.6.3) and 21.2, this becomes

$$\frac{1}{c}\int f(x)\, d\nu(x) = c\int f(x)\, d\nu(x), \tag{21.6.4}$$

whence $c = 1$.

21.7. We observe that the function $d(x)$, defined for a right-invariant measure μ analogously to the definition of $c(x)$ for ν, i.e. defined by $\mu(aM) = d(a)\mu(M)$, bears a very close relationship to $c(x)$. For if we take the right-invariant measure $\mu(M) = \nu(M^{-1})$, where ν is left-invariant, we obtain

$$\mu(aM) = \nu(M^{-1}a^{-1}) = c(a^{-1})\nu(M^{-1}) = c(a^{-1})\mu(M),$$

whence $d(a) = c(a^{-1}) = 1/c(a)$.

21.8. As a final consequence of the uniqueness theorem we make the following remark. In the construction of Haar measure (cf. §15) we chose arbitrarily a prime ideal $\mathfrak{P}$ containing a certain given ideal $\mathcal{T}$ and we defined Haar measure as the $\mathfrak{P}$-limit of certain sequences. Knowing the uniqueness theorem enables us to assert that we should have obtained the same limit regardless of our choice of $\mathfrak{P}$ at least for a very wide class of sets C. (Cf. §15, in particular 15.5.)

More precisely:

21.8.1. *The compact set C has the boundary $C \cdot \widetilde{C^i}$. This boundary is of measure 0, i.e. $\nu(C \cdot \widetilde{C^i}) = 0$, if and only if $\nu(C) = \nu(C^i)$.*

21.8.2. *Let C, E, A be compact sets, C, E having boundaries of measure 0 (cf. above), and $E^i, A^i \neq \theta$. Form the ideal $\mathcal{T}$ of 15.4.2.1 and 15.4.2.2. Then the*

$$\lambda_A(C) = \frac{n[\frac{C}{A}]}{n[\frac{E}{A}]}$$

of (15.1.3) converge to $\frac{\nu(C)}{\nu(E)}$ modulo $\mathcal{T}$.

PROOF. Ad 21.8.1: Obvious, remembering that C is closed by 5.1, and that $\nu(C)$ is finite.

Ad 21.8.2: Consider a prime ideal $\mathfrak{P} \geqq \mathcal{T}$ and the (unique) limit function $\varphi_{\mathfrak{P}}(\xi)$ with the independence ideal $\mathfrak{P}$ (cf. 10.5.3 and 15.5). Use this $\varphi_{\mathfrak{P}}(\xi)$ and $\mathfrak{P}$ in 15.5 (for its $\varphi(\xi)$ and $\mathfrak{P}$), and form, as described there,

$$(\#) \qquad \lambda(C) = \varphi_{\mathfrak{P}}(\lambda_A(C)/A \in \mathcal{T})$$

(cf. (15.5.1)). As discussed in 15.5 and 14.1 this $\lambda(C)$ generates a measure $\nu_1(M)$ in the sense of 4.2, which is a Haar measure. (We write $\nu_1(M)$, instead of $\nu(M)$ as loc. cit., in order to distinguish it from the fixed Haar measure $\nu(M)$, which occurs in our statements 21.8.1, 21.8.2.)

$\nu_1(M), \nu(M)$ are both Haar measures; hence the uniqueness theorem 20.8 asserts that

$$(*) \qquad \nu_1(M) = c\nu(M)$$

for a positive finite constant c.

The boundary of C has (ν-)measure 0, so $\nu(C) = \nu(C^i)$ (by 21.8.1), $\nu_1(C) = \nu_1(C^i)$, (by $(*)$). Consequently 2.3.1, (4.1.1) give

$$\text{(i)} \qquad \lambda(C) = \nu_1(C).$$

Replacing, for a moment, C by E gives similarly

$$\text{(ii)} \qquad \lambda(E) = \nu_1(E).$$

However it is clear that always $\lambda_a(E) = 1$; hence $\lambda(E) = 1$. Consequently (ii) becomes $\nu_1(E) = 1$, and therefore (i) may be written

$$\text{(iii)} \qquad \lambda(C) = \frac{\nu_1(C)}{\nu_1(E)}.$$

Combining (iii) and $(*)$ gives

$$\text{(iv)} \qquad \lambda(C) = \frac{\nu(C)}{\nu(E)}.$$

Now (iv) transforms (#) into

$$(\#\#) \qquad \varphi_{\mathfrak{P}}(\lambda_A(C)|A \in \mathcal{T}) = \frac{\nu(C)}{\nu(E)}.$$

The right-hand side is independent of the choice of $\mathfrak{P}$—i.e. it is the common value of the left-hand side for all prime ideals $\mathfrak{P} \geqq \mathcal{T}$. Hence by 11.4.4 $\lambda_A(C)$ converges to $\frac{\nu(C)}{\nu(E)}$ modulo $\mathcal{T}$.

Thus the proof is completed.

21.8.3. It seems worthwhile to restate 21.8.2 in an explicit form, i.e. having substituted into it the definitions 11.3 of convergence modulo $\mathcal{T}$ and 15.4 of the ideal $\mathcal{T}$. In this way the following statement is obtained:

Let C, E be compact sets, both with boundaries of measure 0, *and $E^i \neq \theta$. Then there exists for every $\varepsilon > 0$ a compact set $A_0 = A_0(C, E, \varepsilon)$ with $1 \in A_0^i$, such that for every compact set A with $1 \in A^i$ and $A \leqq A_0$*

$$\left| \frac{n[\frac{C}{A}]}{n[\frac{E}{A}]} - \frac{\nu(C)}{\nu(E)} \right| < \varepsilon.$$

21.9. The result 21.8.3 is quite remarkable, even when viewed as a geometrical result concerning the simplest Lie groups. Even for the n-parameter Abelian group (i.e. the addition group of n-dimensional vectors) it is far from trivial, since the sets C, E, A may have very complicated shapes—indeed A may have a boundary of positive measure. And for a general Lie group it conveys a structural insight which is probably entirely unamenable to direct approach.

The heuristic considerations of 14.3 were the basis of our presentation of Haar's construction,—21.8.3 gives them an exact meaning. Continuing this line of thought, however, the following observation suggests itself: Since the quotients

$$\lambda_A(C) = \frac{n[\frac{C}{A}]}{n[\frac{E}{A}]}$$

are shown by 21.8.3 to be convergent in the ordinary sense (when A contracts on 1—cf. the exact formulation loc. cit.) it seems peculiar that a highly transcendental method, involving the generalized limits of Chapter II, had to be applied.

I.e.: Is it inescapable to obtain the existence of Haar's measure by applying these generalized limits to the sequences only to discover later, with the help of this existence and of the uniqueness theorem, that these sequences converge in the ordinary sense! This is a very paradoxical situation, and strongly motivates the suspicion that a *direct proof of existence* for Haar's measure ought to exist. That is, a *finite proof*, which proves the convergence of the sequences $(\lambda_A(C)|A \in \mathcal{T})$ in the ordinary sense, by *explicit evaluations.*

We shall see in Chapter VI that such a direct proof can indeed be found. It will combine some basic components of the uniqueness considerations of this chapter, with certain combinatorial results and procedures of an entirely different character.

CHAPTER V

Measure and topology

22. Preliminary remarks

In previous chapters we have seen that under suitable conditions a topology in a group determines one and only one (left-invariant) measure, to be called the Haar measure, in the group. This being settled, various other problems present themselves, which are in different ways *inverse problems* to the above one, i.e. to the existence of a unique Haar measure, based on a given topology. The main questions of this type are these:

22.1.1. The "suitable condition" on the topology referred to above, which was found to be sufficient for the existence of a Haar measure, was the local compactness of the group. Is this condition necessary? Or may we relax it, and still have a Haar measure?

22.1.2. Can the same group, when provided with two different topologies, have the same Haar measure? Or does the Haar measure, in a topological group, in turn determine the topology uniquely?

22.1.3. In 22.1.2 the group was assumed to be topological. Let us now go a step further, and consider a group with a Haar measure, in which no topology is given. Does the Haar measure by itself make the definition of a topology possible? Will this topology be of necessity unique and locally compact? (The two last questions bring us back to 22.1.2 and 22.1.1, respectively.)

22.2. With the exception of 22.1.2, when applied to a locally compact group, none of the three above problems can be attacked, without a careful reconsideration of what constitutes a Haar measure. This discussion, together with the steps needed to solve all three problems 22.1.1–22.1.3, would lead essentially beyond the formal framework of these discussions, and therefore it will not be done here. We shall answer the questions only of 22.1.2, assuming a locally compact group. But we shall give a qualitative discussion of the interconnectedness of these problems, hoping that this will put the reader into a better position to understand and to evaluate the existing literature.

The entire complex of inverse problems was first formulated and solved by A. Weil, to whose work, which includes a solution of 22.1.3, the reader is hereby referred.[1]

[1]First announcement:

(1) A. Weil, Comptes Rendus.

Detailed exposition:

(2) A. Weil, *L'intégration dans les groupes topologiques*, Actualités Scientifiques et Industrielles, No. 863 (1940), Hermann & Cie, Paris.

See also

(3) K. Kodaira, *Über die Beziehung zwischen den Massen und Topologien in einer Gruppe*, Proc. Phys.-Math. Soc. Japan, 23 (1941) 67–119.

22.3. So far, a Haar measure has been characterized as a nonnegative, completely additive, left-invariant set function which possesses the following further properties:

22.3.1. It is defined for all Borel sets.

22.3.2. It is positive for open sets ($\neq \theta$).

22.3.3. It is finite for compact sets.

22.3.4. It is regular.

All four of the conditions stated here depend on topological notions. Yet 22.1.1—and possibly 22.1.2 as well—deals with a topological group, in which the requirements concerning topology have been indefinitely relaxed. And 22.1.3 deals with one in which no topology is given at all. What, under such conditions, are we to understand by a Haar measure?

Thus we must attempt to replace 22.3.1–22.3.4 by other conditions which involve measure alone, without any reference to a topology. The only feasible procedure to find such conditions is to try to describe the fundamental notions of topology in terms of the Haar measure in the case with which we are already familiar: in a locally compact topological group. If this description is successfully carried out we may hope to generalize it so as to give the necessary conditions under which the inverse problems—mainly 22.1.3, which is clearly the most important one—can be solved.

For some aspects of 22.3.1–22.3.4 this description is easy, with no more information than we possess already. Thus 22.3.1 ought to be replaced by the requirement that the measure (i.e. the set function in question) must be defined on a Borel field of sets (cf. 17.4). In locally compact groups 22.3.2 is easily seen to be equivalent to the assertion that the measure is not identically zero.[2] In locally compact groups, in the presence of regularity, 22.3.3 is equivalent to the requirement that the measure be not identically infinite for sets $\neq \theta$.[3] Thus the decisive difficulty is presented by the remaining condition 22.3.4: the notion of regularity. This notion was indeed defined in §18 on an entirely topological basis. What intrinsic—untopological—property of the measure does it represent?

Our main use of regularity was the proof of Fubini's theorem in §20. And that theorem was the most important tool in proving the uniqueness of Haar's measure in §21. Besides it is an intrinsic—untopological—property of measure.

All these reasons make it plausible that Fubini's theorem—or something like it—will be the looked for substitute of regularity, i.e. of the condition 22.3.4. The results of A. Weil (cf. footnote 1, p. 71) justify this view.

A more immediate understanding of the intrinsic measure theoretical meaning of the main notions of topology will be achieved in §§24, 25.

22.4. The heuristic discussion of 22.3 indicates that we must first express the main notions of topology in terms of Haar measure, in those cases where our theory already applies, i.e. for locally compact groups. And this is also the obvious way to settle the questions of 22.1.2 in the sense outlined at its end.

---

[2] Assume O and $\neq \theta, \nu(O) = 0$. Choose $a \in O$; replacing O by $a^{-1}O$ makes $a = 1$. For any compact C we have $C \leqq \sum_{b \in c} bO$; hence $C \leqq \sum_{i=1}^{n} b_i O$.

Thus $\nu(O) = 0$ implies $\nu(C) = 0$. Now the definition of the domain $\mathcal{B}(\mathcal{R})$ of $\nu(M)$, as discussed at the beginning of §18, shows that always $\nu(M) = 0$.

[3] The argument of 18.7 shows that 22.3.3 is equivalent to every point having a finite measure. By left invariance all points have the same measure. Hence the negation of 22.3.3 means that all points, and consequently all sets $\neq \theta$, have infinite measure.

We shall succeed in doing this as a byproduct of some general results on the connection between measure theory and topology, to which the remaining sections of this chapter are devoted.

For the solution of 22.1.1 and 22.1.3 we refer, as stated above, to the work of A. Weil (cf. footnote 1, p. 71). It turns out that local compactness can be replaced by the weaker notion of *local total boundedness*, but that every locally totally bounded group can be extended uniquely, in a very simple and natural way, to a locally compact one. The furthergoing properties of these locally totally bounded groups, which can occur in this connection, should be investigated further. The necessary considerations fit naturally into the framework of a general, purely combinatorial—i.e. untopological—measure theory. We hope to undertake them exhaustively on another occasion.

23. Hilbert space

Throughout this section and the three following, we assume that G is a locally compact group, and $\nu(M)$ a Haar measure in G,—i.e. left-invariant and regular, defined for the Borel sets in G.

We propose to discuss the Hilbert space $\mathfrak{H}$ of all complex-valued functions $f(x)$, $x \in G$, for which

$$\int |f(x)|^2 \, d\nu(x) \text{ is finite.}$$

We must, of course, restrict ourselves to some class of functions in which the above integral can always be formed. This is best achieved with the means at our disposal, by demanding that $f(x)$ be a Baire function in the sense of 19.1. Since 19.1 was, however, restricted to functions $f(x)$ which are everywhere (real and) $\geqq 0$, we must add the following explanation:

23.1.1. *A real-valued function* $f(x)$ *is a Baire function, if the two functions*

$$\text{Max}(f(x), 0), \quad \text{Max}(-f(x), 0)$$

are both Baire functions in the sense of 19.1.

(Those two functions are always $\geqq 0$, and 19.1 applied to such functions only.)

Remembering 19.1, this means obviously: For every $\alpha > 0$ the act of all x with $f(x) > \alpha$ as well as the set of all x with $f(x) < \alpha$ is a Borel set.

23.1.2. *A complex-valued function* $f(x)$ *is a Baire function, if the two functions*

$$\text{Re}\, f(x), \quad \text{Im}\, f(x)$$

are both Baire functions in the sense of 23.1.1.

The remarks following 19.1 may be repeated: It is well known that this family of functions is closed under the operations of addition, subtraction, multiplication, and passage to limit (of an everywhere convergent sequence), and that a theory of integration, analogous in all respects to the usual Lebesgue theory, can be developed in it for any measure of the type we are considering, and we shall freely make use, in what follows, of the concepts of integral, integrability of a Baire function, etc. (Since the functions under consideration are no longer necessarily $\geqq 0$, only finite integrals are now admissible—this one point differing from 19.1.)

Of course all these considerations can be easily based on the corresponding ones mentioned in 19.1, by trivial operations of linear combination.

We can now formulate the exact definition of $\mathfrak{H}$:

23.1.3. The Hilbert space of G to be denoted by $\mathfrak{H} = \mathfrak{H}(G)$ is the set of all complex-valued Baire functions $f(x)$, $x \in G$ for which

$$\int |f(x)|^2 \, d\nu(x) \text{ is finite.}$$

Two such functions, when differing only on an x-set of measure zero, are considered as representing the same element of $\mathfrak{H}$.

It is well known that for any two $f(x), g(x)$ of $\mathfrak{H}$ the functions $\alpha f(x)$ (α any complex constant), $f(x) + g(x)$ belong also to $\mathfrak{H}$. Also the integral

$$\int f(x)\overline{g(x)} \, d\nu(x) \text{ exists.}$$

(I.e. the integrand is absolutely summable, in the usual sense of Lebesgue integration.) We write, as usual,

$$(f, g) = \int f(x)\overline{g(x)} \, d\nu(x),$$

and

$$|f| = \sqrt{(f, f)} = \sqrt{\int |f(x)|^2 \, d\nu(x)}.$$

For these facts, as well as for the elementary geometrical properties of Hilbert space, and of its fundamental operations

$$\alpha f, \quad f + g, \quad (f, g), \quad |f|,$$

cf. any standard treatise on the subject.[4]

---

[4] We mention the following:

[1] J. von Neumann, *Allgemeine Eigenwerttheorie Hermitescher Funktional operatoren*, Math. Annalen, vol. 102 (1929), pp. 49–131.

[2] M. H. Stone, *Linear transformations in Hilbert space*, Amer. Math. Soc. Colloquium Publications Series, vol. XV (1932), New York.

[3] J. von Neumann, *Functional operators*, mimeographed Princeton Mathematical Notes for the years 1934–35 (Part I) and 1935–36 (Part II), Princeton.

[4] F. J. Murray, *Linear transformations in Hilbert space*, Annals of Mathematics Studies, No. 4 (1941), Princeton.

The pertinent parts are: [1], pp. 63–78, 108–111, or [2], pp. 1–32 (pp. 33–35 are also of interest), or [3], Part II, Chap. XII, or [4], pp. 4–21, 27–30. [4] is the most recent exposition, and gives an excellent concise account of the present state of the theory.

The reader might experience a certain difficulty due to the fact that we have to use the term "Hilbert space" in a natural sense, i.e. without necessarily implying either infinite dimensionality or separability of the space $\mathfrak{H}$. I.e., of the five postulates A–E which characterize a Hilbert space (cf. [1], pp. 64–66, or [2], pp. 3–4, or [3], Part II, Chap. XII, p. 2.15, or [4], pp. 4–6), we assume only A,B, and E.

This does not affect the discussion of [3], because there these conditions receive full consideration. (Loc. cit., the postulates C and D are replaced by disjunctions $C_1 - C_2$ and $D_1 - D_2$, between which we need not choose.) The nonseparable case was discovered independently by H. Löwig, F. Rellich, Y. Y. Tseng, and J. von Neumann. In [1], [2], [4] the corresponding discussions must, however, be taken cum grano salis, since they formally presuppose C and D. In fact these considerations which we need here (equally in [1], [2], and [4]), do not make use of C and D (often even not of E), as the reader can easily verify himself. If these circumstances are kept in mind, the reader will be able to refer to any one of the treatises [1]–[4] without trouble.

The considerations which will be mainly used here are these:

23.2. For any set M ($\leqq G$) we denote by χ_M the *characteristic function* of M:

$$\chi_M(x)\begin{cases} = 1 & \text{for } x \in M, \\ = 0 & \text{for } x \notin M. \end{cases}$$

One verifies immediately:

23.2.1. *χ_M belongs to $\mathfrak{H}$ if and only if M is a Borel set and $\nu(M)$ is finite.*

23.2.2. *If χ_M, χ_N belong to $\mathfrak{H}$ (cf. 23.2.1), then*

$$(\chi_M, \chi_N) = \nu(M \cdot N),$$
$$|\chi_M| = \sqrt{\nu(M)},$$
$$|\chi_M - \chi_N| = \sqrt{\nu((M+N)\cdot(\widetilde{M \cdot N}))}$$

We now prove

23.2.3. *The χ_C (C any compact set $\leqq G$) all belong to $\mathfrak{H}$ and the smallest closed linear set containing them all is $\mathfrak{H}$.*

PROOF. Since all $\nu(C)$ are finite, all χ_C belong to $\mathfrak{H}$ by 23.2.1.

The finite linear combinations of all χ_M in $\mathfrak{H}$ (cf. 23.2.1) are dense in $\mathfrak{H}$ (cf. (δ) at the end of footnote 4, below). I.e. the smallest closed linear set containing all χ_M in $\mathfrak{H}$ is $\mathfrak{H}$. In order to prove the same thing for the totality of all χ_C it suffices therefore to prove that the χ_C are dense in the set of all χ_M in $\mathfrak{H}$. I.e.: Given a χ_M in $\mathfrak{H}$ and an $\varepsilon > 0$, we must find a compact C with $|\chi_M - \chi_C| < \varepsilon$.

Now $\nu(M)$ is finite by 23.2.1; hence regularity implies the existence of a $C \leqq M$ with $\nu(C) > \nu(M) = \varepsilon^2$ by 18.1. So $\nu(M \cdot \widetilde{C}) < \varepsilon^2$. But $C \leqq M$ and 23.2.2 give $|\chi_M - \chi_C| = \sqrt{\nu(M \cdot \widetilde{C})}$; consequently $|\chi_M - \chi_C| < \varepsilon$, as desired.

23.3. Owing to the left invariance of $\nu(M)$ it is clear that for any fixed $a \in G$, $f(ax)$ belongs to $\mathfrak{H}$ whenever $f(x)$ does. Also, that $f(ax)$ undergoes only a change on an x-set of measure zero, if $f(x)$ is so changed. So we can define an operator U_a in $\mathfrak{H}$ (i.e. a function with the domain $\mathfrak{H}$ and the range in $\mathfrak{H}$) by

$$U_a f(x) = f(ax).$$

This U_a is clearly everywhere defined in $\mathfrak{H}$ and linear.

The left invariance of $\nu(M)$ also gives

$$\int |f(ax)|^2 \, d\nu(x) = \int |(f(x))|^2 \, d\nu(x),$$

(α) Schwarz's inequality: [1], p. 64, or [2], p. 5, or [3], Part II, Chapter XII, p. 2.3, or [4], pp. 5–6.

(β) The metric of Hilbert space: [1], p. 65, or [2], pp. 5–6, or [3], loc. cit. p. 2.4, or [4], pp. 6–7.

(γ) Riesz's lemma on linear functionals: [1], p. 94 (footnote 52)), or [2], pp. 62–63, or [4], pp. 12–13.

(δ) Approximation of any f ($\in \mathfrak{H}$) by finite linear aggregates of characteristic functions (cf. the beginning of the proof of 23.2.3 below). This is explicitly stated, free of any assumption of separability, in [3], Part II, Chap. XII, p. 2.32.4 (proof of Theorem 12′.6, discussion of the sets S_1, S_2, S_3). The corresponding considerations in [1], [2] and [4] are: [1], middle of p. 110, or [2], pp. 24–25, or [4], p. 28. In all these cases separability was assumed, but can be disregarded.

i.e.

$$|U_a f| = |f|.$$

Summing up:

23.3.1. *The above operator U_a is unitary (cf. Koopman, Proc. N.A.S. 1931, p. 315).*

One also verifies immediately:

23.3.2. *The U_a, $a \in G$, form a representation of the dual of (cf. 14.2.1), i.e.*

$$U_1 = I, \qquad U_b U_a = U_{ab}.$$

After these preliminaries we prove

23.3.3. *For every F, $U_a f$ is a uniformly continuous ($\mathfrak{H}$-valued) function of a in G. I.e.: Given f and any $\varepsilon > 0$, there exists an open $O_1 = O_1(f, \varepsilon)$, such that $1 \in O_1$ ($\leqq G$), and that $ab^{-1} \in O_1$ implies*

$$|U_a f - U_b f| < \varepsilon.$$

23.3.4. *For every f, g $(U_a f, g)$ is a uniformly continuous (numerical, complex-valued) function of a in G. I.e.: Given f, g and any $\varepsilon > 0$, there exists an open $O_2 = O_2(f, g, \varepsilon)$ such that $1 \in O_2$ ($\leqq G$), and that $ab^{-1} \in O_2$ implies*

$$|(U_a f, g) - (U_b f, g)| < \varepsilon.$$

PROOF. By Schwarz's inequality (cf. (α) at the end of footnote 4, p. 75)

$$|(U_a f, g) - (U_b f, g)| = |(U_a f - U_b f, g)| \leqq |U_a f - U_b f| \cdot |g|,$$

so that 23.3.4 is a consequence of 23.3.3. (Simply replace $\varepsilon > 0$ by an $\varepsilon' > 0$ with $\varepsilon' \cdot |g| \leqq \varepsilon$). As to 23.3.3 we observe that by 23.3.1 and 23.3.2

$$|U_a f - U_b f| = |U_b U_{ab}^{-1} f - U_b f| = |U_b(U_{ab^{-1}} f - f)| = |U_{ab^{-1}} f - f|,$$

so that we can replace a, b by $ab^{-1}, 1$. I.e., we may assume further that $b = 1$.

Ad 23.3.3 with $b = 1$: By 23.3.1

$$|U_a g - U_a h| = |U_a(g - h)| = |g - h|;$$

hence $U_a g$ is (for all $a \in G$) a uniformly continuous function of g. Consequently the set of all those f for which 23.3.3 with $b = 1$ is true is closed. It is obviously linear too. Hence it suffices, by 23.2.3, to prove that all χ_C, C compact, belong to this set. I.e.: Given a compact C and any $\varepsilon > 0$, find an open $O_1' = O_1'(C, \varepsilon)$ such that $I \in O_1'$ ($\leqq G$) and that $a \in O_1'$ implies

$$|U_a \chi_C - \chi_C| < \varepsilon.$$

Since $U_a \chi_C = \chi_{a^{-1}C}$ and owing to 23.2.2, this amounts to

$$\nu((C + a^{-1}C) \cdot (\widetilde{C \cdot a^{-1}C})) < \varepsilon^2.$$

Now $(C + a^{-1}C) \cdot (\widetilde{C \cdot a^{-1}C}) = C \cdot \widetilde{a^{-1}C} + a^{-1}C \cdot \widetilde{C}$ (disjointly), so

$$\nu((C + a^{-1}C)(\widetilde{C \cdot a^{-1}C})) = \nu(C \cdot \widetilde{a^{-1}C}) + \nu(a^{-1}C \cdot \widetilde{C})$$

and

$$\nu(C \cdot \widetilde{a^{-1}C}) = \nu(C) - \nu(\widetilde{C} \cdot \widetilde{a^{-1}C}),$$

$$\nu(a^{-1}C \cdot \widetilde{C}) = \nu(a^{-1}C) - \nu(\widetilde{C} \cdot \widetilde{a^{-1}C}),$$

hence $\nu(C) = \nu(a^{-1}C)$ gives

$$\nu(C \cdot \widetilde{a^{-1}C}) = \nu(a^{-1}C \cdot \widetilde{C}),$$

and so

$$\nu((C \cdot a^{-1}C)(\widetilde{C \cdot a^{-1}C})) = 2\nu(a^{-1}C \cdot \widetilde{C}).$$

So we must establish that

$$\nu(a^{-1}C \cdot \widetilde{C}) < \frac{1}{2}\varepsilon^2.$$

Regularity implies the existence of an open $O \geqq C$ with $\nu(O) < \nu(C) + \frac{1}{2}\varepsilon^2$, by 18.2. Owing to the continuity of $a^{-1}x$ (as a two-variable function) in a, x and owing to the compactness of C, there exists an O_1' with $1 \notin O_1'$ such that $a \in O_1'$, $x \in C$ imply $a^{-1}x \in O$. (Cf. for the analogous proof of a similar statement at the end of §8 in Chapter VI.) I.e. $a \in O_1'$ implies $a^{-1}C \leqq O$. But then

$$\nu(a^{-1}C \cdot \widetilde{C}) \leqq \nu(O \cdot \widetilde{C}) = \nu(O) - \nu(C) < \frac{1}{2}\varepsilon^2,$$

as desired.

24. Characterizations of the topology

We are now able to give several characterizations of the topology of G in terms of the measure $\nu(M)$. This carries through the program formulated at the beginning of 22.4; i.e. it answers the questions of 22.1.2, in the affirmative sense of its last part. We formulate all these characterizations together, since they are best proved together.

The topology of G will be characterized by criteria which state when an x ($\in G$) is an *inner point* of a given set A ($\leqq G$). This does indeed determine when A is open, and consequently also when A is closed.

24.1.1. *Given a finite system $f_1, \ldots, f_n \in \mathfrak{H}$, an $\varepsilon > 0$, and an $a \in G$ denote by $N_1^O(f_1, \ldots, f_n; \varepsilon, a)$ the set of all $b \in G$ with*

$$|U_a f_k - U_b f_k| < \varepsilon \quad \textit{for all } k = 1, \ldots, n.$$

Then a is an inner point of A if and only if some $N_1^O(f_1, \ldots, f_n; \varepsilon, a) \leqq A$.

24.1.2. *The same is true if we restrict ourselves to the $N_1^O(f; \varepsilon, a)$ with $n = 1$.*

24.1.3. *Given a finite system $f_1, g_1, \ldots, f_n, g_n \in \mathfrak{H}$, an $\varepsilon > 0$, and an $a \in G$, denote by $N_2^O(f_1, g_1, \ldots, f_n, g_n; \varepsilon, a)$ the set of all $b \in G$ with*

$$|(U_a f_k, g_k) - (U_b f_k, g_k)| < \varepsilon \quad \textit{for all } k = 1, \ldots, n.$$

Then a is an inner point of A if and only if some $N_2^O(f_1, g_1, \ldots, f_n, g_n; \varepsilon, a) \leqq A$.

24.1.4. *The same is true if we restrict ourselves to the $N_2^O(f, g; \varepsilon, a)$ with $n = 1$.*

24.1.5. *For $a = 1$ we may even restrict ourselves to the $N_2^O(f, f; \varepsilon, 1)$ with $n = 1$ and $f = g$. We can also assume that $|f| = 1$.*

24.1.6. *For $a = 1$, $|f| = 1$, the above $N_2^O(f, f; \varepsilon, 1)$ is the set of all $b \in G$ with*

(24.1.6.1) $$|(U_b f, f) - 1| < \varepsilon.$$

We can replace it by the set $N_3^O(f; \varepsilon)$ of all $b \in G$ with

(24.1.6.2) $$|(U_b f, f)| > 1 - \varepsilon,$$

or equally by the set $N_4^O(f;\varepsilon)$ of all $b \in G$ with

(24.1.6.3) $$\mathrm{Re}(U_b f, f) > 1 - \varepsilon.$$

24.1.7. *Given a finite system of Borel sets $M_1, N_1, \dots, M_n, N_n$ with $\nu(M_1)\nu(N_1), \dots, \nu(M_n), \nu(N_n)$ finite, an $\varepsilon > 0$, and an $a \in G$, denote by $M_5^O(M_1, N_1, \dots, M_n, N_n; \varepsilon, a)$ the set of all $b \in G$ with*

$$|\nu(aM_k \cdot N_k) - \nu(bM_k \cdot N_k)| < \varepsilon \quad \textit{for all } k = 1, \dots, n.$$

Then a is an inner point of A if and only if some $N_5^O(M_1, N_1, \dots, M_n, N_n; \varepsilon, a) \leqq A$.

24.1.8. *The same is true if we restrict ourselves to the $N_5^O(M, N; \varepsilon, a)$ with $n = 1$.*

24.1.9. *For $a = 1$ we may even restrict ourselves to the $N_5^O(M, M; \varepsilon, 1)$ with $n = 1$ and $M = N$.*

24.1.10. *For $a = 1$, the above $N_5^O(M, M; \varepsilon, 1)$ is the set of all $b \in G$ with*

$$\nu(bM \cdot M) > \nu(M) - \varepsilon.$$

24.1.11. *Given a Borel set M with $\nu(M) > 0$, denote by $N_6^O(M; a)$ the set aMM^{-1}. Then a is an inner point of A if and only if some $N_6^O(M; n) \leqq A$.*

PROOF. Observe first that wherever a general a occurs we may replace a, A by $1, a^{-1}A$. I.e. we can, and will, assume $a = 1$ throughout what follows.

We must prove two things:

(α) 1 is an inner point of each one of the above sets N^O.

(β) If O is open and $1 \in O$, then there exists a set $N^O \leqq O$ for each one of the above described categories of sets N^O.

Proof of (α). In this case it suffices to prove 24.1.1, 24.1.3; the others are special cases of these. Indeed: 24.1.2 is a special case of 24.1.1; 24.1.4, 24.1.5 are special cases of 24.1.3. In 24.1.6 the formula (24.1.6.1) is evident, since $(U_a f, f) = (f, f) = |f|^2 = 1$. The $N_3^O(f;\varepsilon)$, $N_4^O(f;\varepsilon)$ of (24.1.6.2), (24.1.6.3) may be used since (24.1.6.1) implies (24.1.6.3), and this again implies (24.1.6.2), so that $N_2^O(f, f; \varepsilon, 1) \leq N_4^O(f;\varepsilon) \leqq N_3^O(f;\varepsilon)$. 24.1.7 is a special case of 24.1.3, with $f_1 = \chi_{M_1}, g_1 = \chi_{N_1}, \dots, f_n = \chi_{M_n}$. 24.1.8, 24.1.9 are special cases of 24.1.7. 24.1.10 is obvious since $\nu(bM \cdot M) \leqq \nu(M)$. 24.1.11 is a consequence of 24.1.10 since $b \in N_6^O(M; 1) = MM^{-1}$ clearly means $bM \cdot M \neq \theta$, hence is a consequence of $\nu(bM \cdot M) > 0$, and therefore

$$N_6^O(M; 1) \geqq N_5^O(M, M; \varepsilon, 1) \quad \text{with } \varepsilon = \nu(M).$$

Proof of (α) for 24.1.1, 24.1.3: Clearly $b = a = 1$ belongs to both sets N^O involved, and both are open sets, owing to 23.3.3, 23.3.4, respectively.[5]

Proof of (β). Here we proceed in the converse direction. In this case it suffices to prove 24.1.11; the others are special cases of this. Indeed: 24.1.10 has been settled above under (α). ((α), (β) do not enter there.) 24.1.9, using 24.1.10, is a special case of 24.1.11 owing to the same inclusion

$$N_6^O(M; 1) \geqq N_5^O(M, M; \varepsilon, 1) \quad \text{with } \varepsilon = \nu(M),$$

which was used in (α) for the converse implication. 24.1.8, 24.1.7 are special cases of 24.1.9. The $N_3^O(f;\varepsilon)$ of (24.1.6.2) may be used as a special case of 24.1.9, using 24.1.10, owing to the following argument: $\nu(M) = 0$ would give $N_5^O(M, M; \varepsilon, 1) =$

[5]This is the only really significant step in the entire proof!

G (by 24.1.10); hence $A = G$, and therefore we may assume that $\nu(M) > 0$. Put $f = \chi_M$; then $f \neq 0$. Now $(U_b f, f) = (\chi_{b^{-1}M}, \chi_N) = \nu(b^{-1}M \cdot M) = \nu(M \cdot bM)$ (by 23.2.2); hence $|(U_b f, f)| = \nu(M \cdot bM)$. Therefore replacement of ε by $\varepsilon|f|^2$, and subsequent replacement of f by $\frac{1}{|f|} \cdot f$, carries 24.1.10 into (24.1.6.2). The $N_2^O(f, f; \varepsilon, 1), N_4^O(f; \varepsilon)$ of (24.1.6.1), (24.1.6.3) may be used, owing to the same inclusions $N_2^O(f, f; \varepsilon, 1) \leqq N_4^O(f; \varepsilon) \leqq N_3^O(f; \varepsilon)$, which were used in ($\alpha$) for the converse implications. The correctness of the formula (24.1.6.1) (with respect to 24.1.5) was settled above, under (α). ((α), (β) do not enter here.) Thus 24.1.6, 24.1.5 are completely established. 24.1.3, 24.1.4 are special cases of 24.1.5. 24.1.2 follows from 24.1.4 by Schwarz's inequality:

$$|(U_a f, g) - (U_b f, g)| \leqq |U_a f - U_b f| \cdot |g|.$$

(Cf. the corresponding part of the proof of 23.3.4.) Again, simply replace $\varepsilon > 0$ by an $\varepsilon' > 0$ with $\varepsilon' \cdot |g| \leqq \varepsilon$. 24.1.1 is a special case of 24.1.2.

Proof of (β) for 24.1.11. $a = 1 \in O$. Since xy^{-1} is continuous (as a two-variable function of x, y), there exists an open P with $1 \in P$, so that $x, y \in P$ imply $xy^{-1} \in O$. I.e. $PP^{-1} \leqq O$. P is open and $\neq \theta$ so $\nu(P) > 0$. Thus $M = P$ meets all requirements.

24.2. The reader will perhaps consider the long list 24.1.1–24.1.11 of equivalent characterizations of the topology of G in terms of the measure $\nu(M)$, as pedantic. We hope, however, that on closer inspection he will convince himself that most of these criteria add some new viewpoint to this question.

We wish to say first something in favor of the seemingly unnecessarily complicated criteria 24.1.1, 24.1.3, 24.1.7.

Consider a space $\mathcal{S}$ which is given without a topology, or the topology of which we choose to ignore. Let a family $\mathcal{F}$ of (numerical) functions $\phi(x)$, $x \in \mathcal{S}$, be given in $\mathcal{S}$. We now wish to topologize $\mathcal{S}$ so that all functions $\phi(x)$ of $\mathcal{F}$ become continuous. Then the obvious procedure is this:

24.2.1. *Given a finite system of functions $\phi_1(x), \ldots, \phi_n(x)$ from $\mathcal{F}$, an $\varepsilon > 0$, and an $a \in \mathcal{S}$, denote by $N^*(\phi_1, \ldots, \phi_n; \varepsilon, a)$ the set of all $b \in \mathcal{S}$ with*

$$|\phi_k(a) - \phi_k(b)| < \varepsilon \quad \textit{for all } k = 1, \ldots, n.$$

We topologize $\mathcal{S}$ by declaring that the $N^*(\phi_1, \ldots, \phi_n; \varepsilon, a)$ are the neighborhoods of a. I.e., a is an inner point of A if and only if some $N^*(\phi_1, \ldots, \phi_n; \varepsilon, a) \leqq A$.

Indeed: It is easy to verify that this is the "weakest" topology of $\mathcal{S}$,[6] which makes all $\phi(x)$ of $\mathcal{F}$ continuous. (In this generality we must permit all $n = 1, 2, 3, \ldots$.)

Now 24.1.3 and 24.1.7 obtain by application of this principle, by choosing for $\mathcal{F}$ the families of functions

$$\phi'(x) = (U_x f, g) \tag{24.2.1.1}$$

or

$$\phi''(x) = \nu(xM \cdot N) \tag{24.2.1.2}$$

[6]I.e. the one for which any set A has most inner points. Or equivalently: For which most sets are open. Or just as well: For which most sets are closed.

Editor's note: This is von Neumann's attempt to say "$\mathcal{S}$ has the fewest open sets of any topology, which makes all the φ of $\mathfrak{F}$ continuous."

respectively. And 24.1.1 obtains analogously by using the nonnumerical functions

$$\phi'''(x) = U_x f, \tag{24.2.1.3}$$

with the topology (i.e. the metric) of $\mathfrak{H}$. In fact, we could replace both (24.2.1.1) and (24.2.1.3) by an analogous consideration of the even more abstract, nonnumerical functions

$$\phi''''(x) = U_x, \tag{24.2.1.4}$$

with a suitable topology of the space of all unitary operators in $\mathfrak{H}$. We shall however not go into further details.

We only want to emphasize that 24.2.1 is a very simple and very general method for the topologization of spaces $\mathcal{S}$. It is the abstract core of widely varied procedures: E.g. in the theory of "weak topologies of infinite direct product spaces", in the theory of "abstract almost-periodic functions", etc. And since (24.2.1.1) and (24.2.1.2) are the simplest numerical functions one can think of a group G in which only a measure $\nu(M)$ is given; therefore these formulae with 24.2.1, i.e. the criteria 24.1.3 and 24.1.7, are quite natural ways to topologize G with the help of its measure.

24.3. Concerning the other criteria: 24.1.2 is a simplified form of 24.1.1, and 24.1.4, 24.1.5 are successive simplifications of 24.1.3. 24.1.7 is a specialization of 24.1.3, but it also presents a possibility of expressing the topology of G with the help of $\nu(M)$ directly, without the intermediary of $\mathfrak{H}$. 24.1.8, 24.1.9, 24.1.10 are again successive simplifications of 24.1.7. 24.1.11 finally represents a new principle of topologization, also avoiding $\mathfrak{H}$ and probably the most immediate of all.

(24.1.6.2), (24.1.6.3), 24.1.9, 24.1.11 should also be compared with our subsequent compactness criteria of a similar structure, i.e. with (25.1.3.1), (25.1.3.2), 25.1.5, 25.1.6. In this connection (24.1.6.2), (24.1.6.3) as well as (25.1.3.1), (25.1.3.2) permit a simple geometrical interpretation, which will be discussed in 25.2.

We mention finally that certain other linear spaces, based on G, could be used instead of $\mathfrak{H}$. E.g.: The space $\mathcal{L}_p$ of all $f(x)$ in G

$$[f]^p = \int |f(x)|^p \, d\nu(x) \text{ finite}$$

($1 \leqq p < \infty$), of course $\mathcal{L}_2 = \mathfrak{H}$. Also the space $\mathcal{M}$ of all Borel sets M with a finite $\nu(M)$, with the distance

$$[M, N] = \nu((M + N) \cdot (\widetilde{M \cdot N})).$$

($\mathcal{M}$ is really being used in 24.1.7–24.1.10.) For some purposes where the avoidance of the topology of G is not vital the space $\mathcal{C}$ of all continuous functions $f(x)$ in G is useful.

It seems to us, nevertheless, that $\mathfrak{H}$ is technically most suitable in the majority of cases.

24.4. Since $\nu(M)$ determines the topology of G it also determines all topologically defined notions in G. Nevertheless it is of interest to express some of these directly in terms of $\nu(M)$.

Thus it can be shown that $\mathfrak{H}$ is a finite-dimensional (Euclidean) space if and only if G is a finite group—the number of dimensions of $\mathfrak{H}$ being equal to the number of elements of G. Also, $\mathfrak{H}$ is separable if and only if G is separable. The proofs of these assertions are quite easy, but we shall not discuss them here.

We shall however give an independent discussion of the notion of compactness in the next section.

25. Characterizations of the notion of compactness

More precisely: We shall characterize in terms of the measure $\nu(M)$ those sets A which are subsets of compact sets, i.e. which have a compact closure. If A is closed, this means exactly compactness.

We shall again give several characterizations, all of which except one are best formulated and proved together.

25.1.1. *Given two $f, g \in \mathfrak{H}$ and an $\varepsilon > 0$, denote by $N_1^c(f, g; \varepsilon)$ the set of all $b \in \mathcal{G}$ with*

$$|(U_b f, g)| \geqq \varepsilon.$$

Then A has a compact closure if and only if some $N_1^c(f, g; \varepsilon) \geqq A$.

25.1.2. *The same is true if we restrict ourselves to the $N_1^c(f, f; \varepsilon)$ with $f = g$. We can also assume that $|f| = 1$.*

25.1.3. *For $|f| = 1$ the above $N_1^c(f, f; \varepsilon)$ is the set of all $b \in \mathcal{G}$ with*

$$|(U_b f, f)| \geqq \varepsilon. \tag{25.1.3.1}$$

We can replace it by the set $N_2^c(f, \varepsilon)$ of all $b \in \mathcal{G}$ with

$$|\operatorname{Re}(U_b f, f)| \geqq \varepsilon. \tag{25.1.3.2}$$

25.1.4. *Given two Borel sets $M, N \in \mathfrak{H}$ and an $\varepsilon > 0$, denote by $N_3^c(M, N; \varepsilon)$ the set of all $b \in \mathcal{G}$ with*

$$\nu(bM \cdot N) \geqq \varepsilon.$$

Then A has a compact closure if and only if some $N_3^c(M, n : \varepsilon) \geqq A$.

25.1.5. *The same is true if we restrict ourselves to the $N_3^c(M, M; \varepsilon)$ with $M = N$.*

PROOF. We must prove two things:

(α) Each one of the above sets N^c has a compact closure.

(β) If C is compact, then there exists a set $N^c \geqq C$, for each one of the above described categories of sets N.

Proof of (α). In this case it suffices to prove 25.1.1; the others are special cases of this. Indeed: 25.1.2 is a special case of 25.1.1. In 25.1.3, (25.1.3.1) is an obvious restatement of the definition of $N_1^c(f, f; \varepsilon)$. The $N_2^c(f; \varepsilon)$ of (25.1.3.2) may be used since (25.1.3.2) implies (25.1.3.1) so that $N_1^c(f, f; \varepsilon) \geqq N_2^c(f; \varepsilon)$. 25.1.4 is a special case of 25.1.1, with $f = \chi_M$, $g = \chi_N$. 25.1.5 is a special case of 25.1.4.

Proof of (α) for 25.1.1. Since $U_b f$ is a uniformly continuous function of f for all b (because it is unitary), and owing to Schwarz's inequality (cf. (α) at the end of footnote 4 on p. 75), it suffices to prove this for all $f, g \in \gamma$, if γ is a dense set in $\mathfrak{H}$. By 23.2.3 we may choose for γ the set of all finite linear aggregates of characteristic functions of compact sets: $h = \sum_{k=1}^{n} u_k \chi_{C_k}$. ($u_1, \ldots, u_n$ complex numbers, $C_1, \ldots, C_n$ compact sets). For such an h we have $h(x) = 0$ for $x \notin C$, where $C = \sum_{k=1}^{n} C_k$ is compact. Let D, E be the corresponding compact sets for f, g, respectively. Then $b \notin DE^{-1}$ excludes that $bx \in D$, $x \in E$ for the same x; hence it gives $U_b f(x) \cdot g(x) = f(bx) \cdot g(x) = 0$ for all x. So $(U_b f, g) = 0$, and

$b \notin N_1^C(f,g;\varepsilon)$. Consequently $N_1^C(f,g;\varepsilon) \leqq DE^{-1}$, and DE^{-1} is compact, along with D, E, by 14.4.

Proof of (β). Here we proceed in the converse direction. In this case it suffices to prove 25.1.5; the others are special cases of this. Indeed: 25.1.4 is a special case of 25.1.5. The $N_2^C(f;\varepsilon)$ of (25.1.3.2) may be used as a special case of 25.1.5, owing to the following argument: $\nu(M) = 0$ would give $N_3^{C}(M,M;\varepsilon) = \theta$; hence $A = \theta$, and therefore we may assume $\nu(M) > 0$. Put $f = \chi_M$; then $f \neq 0$. Now $(U_b f, f) = (\chi_{b^{-1}M}, \chi_M) = \nu(b^{-1}M \cdot M) = \nu(M \cdot bM)$ by 23.2.2; hence $\mathcal{R}(U_b f, f) = \nu(M \cdot bM)$. Therefore replacement of ε by $\varepsilon \cdot |f|^2$, and subsequent replacement of f by $\frac{1}{|f|} \cdot f$, carries 25.1.5 into (25.1.3.2). Then $N_1^c(f,f;\varepsilon)$ of (25.1.3.1) may be used, owing to the same inclusion $N_1^c(f,f;\varepsilon) \geqq N_1^c(f;\varepsilon)$, which was used in ($\alpha$) for the converse implication. Formula (25.1.3.1) obviously defines the same set $N_1^c(f,f;\varepsilon)$, as 25.1.2. Thus 25.1.2, 25.1.3 are completely established. 25.1.1 is a special case of 25.1.2.

Proof of (β) for 25.1.5. Let a compact C be given. Choose a compact D with $\nu(D) > 0$.[7] Then $C^{-1}D$ is also compact, by 14.4, and so is $E = C^{-1}D + D$. So $\nu(E)$ is finite. Now $E \geqq D$, and for $b \in C$, $E \geqq C^{-1}D \geqq b^{-1}D$, $bE \geqq D$. Hence $bE \cdot E \geqq D$, $\nu(bE \cdot E) \geqq \nu(D)$. So if we choose $M = E$ and $\varepsilon = \nu(D) > 0$, then $b \in N_3^c(M,M;\varepsilon)$. I.e. $N_3^c(M,M;\varepsilon) \geqq C$, as desired.

25.1.6. *A has a compact closure if and only if there exists a Borel set M with $\nu(M) > 0$ and $\nu(AM)$ finite.*

PROOF. Sufficiency: Let us proceed *á contrario*. Assume that A has no compact closure, i.e. that A is a subset of no compact set. We must prove that then $\nu(M) > 0$ implies $\nu(AM)$ infinite. By 18.1 there exists a compact $D \leqq M$ with $\nu(D) > 0$, so we may replace M by D.

We define an infinite sequence of elements $a_1, a_2, \ldots$ ($\in \mathcal{G}$) as follows: Assume that $a_1, \ldots, a_{i-1}$ ($i = 1, 2 \ldots$) are already defined. DD^{-1} is compact, by 14.4; hence each $a_j DD^{-1}$ and $\sum_{j=1}^{i=1} a_j DD^{-1}$ is too. So $A \nleqq \sum_{j=1}^{i=1} a_j DD^{-1}$. Choose $a_i \in A$ with $a_i \notin \sum_{j=1}^{i=1} a_j DD^{-1}$.

Thus all $a_i \in A$. Hence $a_i D \leqq AD$. Also for $i > j$, $a_i \notin a_j DD^{-1}$, i.e. $a_i D$, $a_j D$ are disjoint. Thus they are, by symmetry, disjoint whenever $i \neq j$. Consequently

$$\nu(AD) \geqq \nu\left(\sum_{i=1}^{\infty} a_i D\right) = \sum_{i=1}^{\infty} \nu(a_i D) = \sum_{i=1}^{\infty} \nu(D) = \infty$$

(since $\nu(D) > 0$). I.e. $\nu(AD)$ is infinite, as desired.

Necessity: Let A have a compact closure, i.e. a compact $C \geqq A$. Choose a compact D with $\nu(D) > 0$. (Cf. footnote 5 on p. 78.) Then CD is compact too by 14.4,. Hence $\nu(CD)$ is finite. Thus $\nu(AD)$ is *a fortiori* finite.

25.2. Reconsidering the criteria 25.1.1–25.1.6 we see that 25.1.1, 25.1.2, (25.1.3.1), (25.1.3.2), 25.1.4, 25.1.5 bear some similarity to 24.1.4, 24.1.5, (24.1.6.2), (24.1.6.3), 24.1.8, 24.1.9, while 25.1.6 may be best compared with 24.1.11. Clearly 25.1.4 is a specialization of 25.1.1 and all other, except 25.1.6, simplified forms of these two. And 25.1.6 seems to be the most immediate formulation of all.

---

[7]E.g. the closure of an open set $O \neq \theta$, which has a compact closure.

There is, however, a geometrical remark which can be made in connection with (24.1.6.2), (24.1.6.3) for open sets, and with (25.1.3.1), (25.1.3.2) for compact sets, and which seems to be quite illuminating.

The *inner product* (f, g) of a Hilbert space is, as is well known, intimately connected with the notion of the *angle*. Indeed, let us normalize f, g by $|f| = |g| = 1$, and denote the angle of the two *directions* f, g by θ. Then there are two plausible definitions for θ: Either

(25.2.1) $$\cos\theta = |(f, g)|,$$ [8]

or

(25.2.2) $$\cos\theta = \mathcal{R}(f, g).$$ [9]

According to which of these two definitions we choose, (24.1.6.2), (24.1.6.3) express that the angle of $U_b f$, f is nearly 0, while (25.1.3.1), (25.1.3.2) express that the angle in question is not nearly $\frac{\pi}{2}$.

26. The density theorem

We conclude this chapter by proving an analogue in topological groups of the Lusin density theorem.

Consider two $f, g \in \mathfrak{H}$ and an open O with $1 \in O$, and compact closure. So $\nu(O)$ is > 0 and finite. Since $(U_a f, g)$ is a continuous function of a, by 23.3.3 we can form the expression

(26.1.1) $$L(f, g) = \frac{1}{\nu(O)} \int_O (U_a f, g)\, d\nu(a).$$

And since by Schwarz's inequality and the unitarity of U_a

$$|(U_a f, g)| \leqq |U_a f| \cdot |g| = |f| \cdot |g|,$$

therefore

(26.1.2) $$|L(f, g)| \leqq |f| \cdot |g|.$$

Obviously $L(f, g)$ is conjugate linear with respect to g; hence (26.1.2) permits us to apply Riesz's lemma on linear functionals. (Cf. (γ) at the end of footnote 4, p. 75.) I.e., there exists a unique $f^* \in \mathfrak{H}$ depending on f, such that for all g

$$L(f, g) = (f^*, g).$$

We write $f^* = A_0 f$; then this becomes

(26.1.3) $$L(f, g) = (A_0 f, g).$$

A_0 is obviously a linear operator, defined for all $f \in \mathfrak{H}$.

[8] In this case the direction of f is the set of all uf, u any complex number; and similarly for g. I.e. a direction is an entire complex line. θ is then normalized by $0 \leqq \theta \leqq \frac{\pi}{2}$.

[9] In this case the direction of f is the set of all uf, u any real number > 0, and similarly for g. I.e. a direction is a real half line. θ is then normalized by $0 \leqq \theta \leqq \pi$.

26.2. We now use the continuity of $U_a f$, as established by 23.3.3, once more. Given an $\varepsilon > 0$, there exists an $O_1 = O_1(f, \varepsilon)$ such that $a \in O$ implies $|f - U_a f| \leqq \varepsilon$. Hence by Schwarz's inequality (cf. also the proof of 23.3.3, 23.3.4),

$$|(f, g) - (U_a f, g)| \leqq \varepsilon \cdot |g|.$$

Assume $O \leqq O_1$; then the above inequality holds for all $a \in O$. So (26.1.1) gives

$$|(f, g) - L(f, g)| \leqq \varepsilon \cdot |g|,$$

and then (26.1.3) gives

$$|(f, g) - (A_O f, g)| \leqq \varepsilon \cdot |g|,$$
$$|(f - A_O f, g)| \leqq \varepsilon \cdot |g|.$$

Put $g = f - A_O f$; this yields immediately

$$|f - A_O f| \leqq \varepsilon. \tag{26.2.1}$$

Consider now a Borel set M with $\nu(M)$ finite. Put $f = \chi_M$; then $O_1 = O_1(f, \varepsilon) = O_1'(M, \varepsilon)$. Before applying (26.2.1) we compute $A_O \chi_M$. (26.1.1), (26.1.3) give:

$$\begin{aligned}
(A_O \chi_M, g) = \mathcal{L}(\chi_M, g) &= \frac{1}{\nu(O)} \int_O (U_a \chi_M, g)\, d\nu(a) \\
&= \frac{1}{\nu(O)} \int_O (\chi_{a^{-1}M} g)\, d\nu(a) \\
&= \frac{1}{\nu(O)} \int_O \left\{ \int_{\mathcal{G}} \chi_{a^{-1}M}(x) \overline{g(x)}\, d\nu(x) \right\} d\nu(a)
\end{aligned}$$

(using Fubini's theorem)

$$\begin{aligned}
&= \frac{1}{\nu(O)} \int_{\mathcal{I}} \left\{ \int_O \chi_{a^{-1}M}(x) \overline{g(x)}\, d\nu(a) \right\} d\nu(x) \\
&= \int_{\mathcal{I}} \frac{1}{\nu(O)} \left\{ \int_O \chi_{a^{-1}M}(x)\, d\nu(a) \right\} \overline{g(x)}\, d\nu(x) \\
&= \int_{\mathcal{I}} \frac{1}{\nu(O)} \left\{ \int_O \chi_{Mx^{-1}}(a)\, d\nu(a) \right\} \overline{g(x)}\, d\nu(x)
\end{aligned}$$

(using 21.1)

$$\begin{aligned}
&= \int_{\mathcal{G}} \frac{\nu(O \cdot Mx^{-1})}{\nu(O)} \overline{g(x)}\, d\nu(x) \\
&= \int_{\mathcal{G}} \frac{\nu(Ox \cdot M)}{\nu(Ox)} \overline{g(x)}\, d\nu(x).
\end{aligned}$$

Consequently

$$A_O \chi_M(x) = \frac{\nu(Ox \cdot M)}{\nu(Ox)}. \tag{26.2.2}$$

And so (26.2.1), (26.2.2) give together:

$$\int \left| \chi_M(x) - \frac{\nu(Ox \cdot M)}{\nu(Ox)} \right|^2 d\nu(x) \leqq \varepsilon. \tag{26.2.3}$$

Given any two $\delta_1, \eta > 0$, put $\varepsilon = \delta^2\eta$, and $O_1 = O_1(f,\varepsilon) = O_1'(M,\delta^2\eta) = O_1''(M,\delta,\eta)$. Then we conclude from (26.2.3):

The set of all x with

$$\left|\chi_M(x) - \frac{\nu(Ox\cdot M)}{\nu(Ox)}\right| \geq \delta \text{ has a measure } \leqq \eta. \tag{26.2.4}$$

Summing up:

26.3. *If M is a Borel set with $\nu(M)$ finite, then there exists for any two δ, $\eta > 0$ an open set $O_1 = O_1''(M,\delta,\eta)$ with $1 \in O_1$, which possesses the following property:*

If O is an open set with $1 \in O$ and with compact closure, and if $O \leqq O_1$, then the set of all x with

$$\frac{\nu(Ox\cdot M)}{\nu(Ox)} \begin{cases} \leqq 1-\delta & \textit{when } x \in M,^{10} \\ \geqq \delta & \textit{when } x \notin M \end{cases}$$

has measure η.

In other words (remembering $0 \leqq \frac{\nu(Ox\cdot M)}{\nu(Ox)} \leqq 1$), $\frac{\nu(Ox\cdot M)}{\nu(Ox)}$ converses *en mesure* to $\chi_M(x)$, i.e. to

$$\begin{cases} 1 & \text{when } x \in M, \\ 0 & \text{when } x \notin M, \end{cases}$$

when O contracts on 1. (For the meaning of the last expression, cf. the corresponding discussion in 21.9.) This limit of $\frac{\nu(Ox\cdot M)}{\nu(Ox)}$ when it exists, may be considered the *density* of M at x. Hence 26.3 may be called the *density theorem*.

It is remarkable that the use of a left-invariant Haar measure $\nu(M)$ caused the appearance of the multiplier x on the right-hand side of O in Ox.

---

[10]This is the same thing as the

$$\left|\chi_M(x) - \frac{\nu(O_x\cdot M)}{\nu(O_x)}\right| \geqq \delta$$

of (26.2.4), since $0 \leqq \frac{\nu(O_x\cdot M)}{\nu(O_x)} \leqq 1$, and

$$\chi_M(x)\begin{cases} = 1 & \text{when } x \in M, \\ = 0 & \text{when } x \notin M. \end{cases}$$

CHAPTER VI

Construction of Haar's invariant measure in groups by approximately equidistributed finite point sets and explicit evaluations of approximations

1. Notations (combinatorics and set theory)

θ : Empty set

$\dot{+} \dot{\sum}$: Set theoretical sum

$+, \sum$: Set theoretical sum of disjoint sets. (Use of these symbols implies assertion of disjointness.)

$\leqq$: Subset relation

$\subset$: Proper subset relation

F, G, H, K, L : Finite set

$F_{(p)}, \ldots, L_{(p)}$: Finite set with (precisely) p elements ($p = 0, 1, 2, \ldots$).

All other capitals : Set. (General, or subject to other restrictions.)

2. Lemma of Hall, Maak and Kakutani[1]

Assume

(a) $A = \sum_{i=1}^{m} A_i$, $B = \sum_{j=1}^{n} B_j$ $(m, n = 0, 1, 2, \ldots)$,

also that

(b) *never* $\sum_{i \in F_{(s)}} A_i \geqq \sum_{j \in G_{(s+1)}} B_j$, *and*

(c) *never* $\sum_{j \in G_{(s)}} B_j \geqq \sum_{i \in F_{(s+1)}} A_i$.

[1] This lemma is a very far reaching generalization of a result of D. König concerning "graphs". Cf. D. König, *Über Graphen und ihre Anwendungen*, Math. Ann., vol. 77 (1916), p. 453. The form in which we made use of it is due to S. Kakutani to whom we are indebted for communicating his result. The problem is also treated with varying degrees of generality in the following papers.

B. L. van der Waerden, *Ein Satz über Klasseneinteilungen von endlichen Mengen*, Hamb. Abh., vol. 5 (1927), p. 185. E. Sperner, *Note zu der Arbeit von Herrn B. L. van der Waerden; 'Ein Satz über Klasseneinteilungen von endlichen Mengen'*, Hamb. Abh., vol. 5 (1927), p. 232.

R. Rado, *Bemerkungen zur Kombinatorik im Anschluss an Untersuchungen von Herrn D. König*, Berl. Berichte, vol. 32 (1933), p. 68. P. Hall, *On representatives of subsets*, J. London Math. Soc., vol. 10 (1935), p. 26.

W. Maak, *Eine neue Definition der fastperiodischen Funktionen*, Hamb. Abh., vol. 11 (1935), p. 240.

R. Rado, *A theorem on general measure functions*, Proc. London Math. Soc., vol. 44 (1938), p. 61.

Then there exist a permutation $(1', \ldots, m')$ *of* $(1, \ldots, m)$ *and a permutation* $(1'', \ldots, n'')$ *of* $(1, \ldots, n)$ *together with a* $p = 0, 1, \ldots, \mathrm{Min}(m, n)$, *such that*

(α) $A_{k'} \cdot B_{k''} \neq \theta$ *for* $k = 1, \ldots, p$,

(β) *whenever* $A_i \leqq B$, *then* $i = k'$ *for a* $k = 1, \ldots, p$,

(γ) *whenever* $B_j \leqq A$, *then* $j = k''$ *for a* $k = 1, \ldots, p$.

PROOF. Assume that this lemma is not always true. Consider a case where it fails, and choose it with its $m + n$ as small as possible.

We first make two observations:

(I) Always $A_i \neq \theta$ and $B_j \neq \theta$. Indeed: Assume the opposite. Assume, by symmetry, that an $A_i = \theta$, say $A_1 = \theta$. Then $\sum_{i \in (1)} A_i = \theta = \sum_{j \in \theta} B_j$, contradicting (c).

(II) At least once $A_i \cdot B_j \neq \theta$. Indeed: Assume the opposite; that is, that always $A_i \cdot B_j = \theta$. Then $A \cdot B = \theta$. Now $A_i \leqq B$ would imply, since $A_i \leqq A$, that $A_i \leqq A \cdot B = \theta$, $A_i = \theta$ contradicting (I). Thus (β) never occurs. Symmetrically (γ) never occurs either. Thus (α)–(γ) could be satisfied with $p = 0$ (and, say, $i' = i$, $j'' = j$). This contradicts our original assumption.

After these preliminaries we proceed to investigate our system more thoroughly. Consider the following assertion:

$$(*) \qquad \begin{cases} \text{If } s = 1, 2, \ldots, \mathrm{Min}(m, n) \text{ but not } s = m = n, \text{ then} \\ \text{(b*) never } \sum_{i \in F_{(s)}} A_i \geqq \sum_{j \in G_{(s)}} B_j \text{ and} \\ \text{(c*) never } \sum_{j \in G_{(s)}} B_j \geqq \sum_{i \in F_{(s)}} A_i. \end{cases}$$

We distinguish two alternatives:

(A) $(*)$ is not always true,

(B) $(*)$ is always true.

For (A): Consider a case where $(*)$ fails. We may assume, by symmetry, that in this case (b*) avails, and, by permuting $(1, \ldots, m)$ and $(1, \ldots, n)$ appropriately, that $F_{(s)} = (1, \ldots, s)$ and that $G_{(s)} = (1, \ldots, s)$. So we have in this case (of the minimum s)

$$(1) \qquad \sum_{i=1}^{s} A_i \geqq \sum_{j=1}^{s} B_j.$$

Now put

$$(2) \qquad A_i' = A_i \left(- \sum_{j=s+1}^{n} B_j \right) \quad \text{for } i = 1, \ldots, s,$$

$$(3) \qquad B_j' = B_j \left(- \sum_{i=1}^{s} A_i \right) \quad \text{for } j = s+1, \ldots, n.$$

We form the two systems

$$(4) \qquad A^1 = \sum_{i=1}^{s} A_i^1, \qquad B^1 = \sum_{j=1}^{s} B_j$$

and

$$(5) \qquad A^2 = \sum_{i=s+1}^{m} A_i, \qquad B^2 = \sum_{j=s+1}^{n} B'_j.$$

We claim that they both satisfy the hypotheses (b), (c) of our lemma. Indeed, let us analyze the opposite possibilities:

(4) violates (b): $\sum_{i\in F_{(t)}} A'_i \geqq \sum_{j\in G_{(t+1)}} B_j$ $(F_{(t)}, G_{(t+1)} \leqq (1,\dots,s))$. Then, *a fortiori*, $\sum_{i\in F_{(t)}} A_i \geqq \sum_{j\in G_{(t+1)}} B_j$, contradicting (b) for the original system.

(4) violates (c): $\sum_{i\in F_{(t+1)}} A'_i \leqq \sum_{j\in G_{(t)}} B_j$ $(F_{(t+1)}, G_{(t)} \leqq (1,\dots,s))$. Subtract this from (1) above; that gives

$$\sum_{i\in F_{(s-t-1)}} A'_i \geqq \sum_{j\in G'_{(s-t)}} B_j, \text{ where } F'_{(s-t-1)} = (1,\dots,s) - F_{(t+1)},$$
$$G'_{(s-t)} = (1,\dots,s) - G_{(t)}.$$

Thus we are back to the preceding case.

(5) violates (c): $\sum_{j\in G_{(t)}} B'_j \geqq \sum_{i\in F_{(t+1)}} A_i$ $(F_{(t+1)} \leqq (s+1,\dots,m))$, $G_{(t)} \leqq (s+1,\dots,n)$. Then, *a fortiori*, $\sum_{j\in G_{(t)}} B_j \geqq \sum_{j\in F_{(t+1)}} A_i$, contradicting (c) for the original system.

(5) violates (b): $\sum_{j\in G_{(t+1)}} B'_j \leqq \sum_{i\in F_{(t)}} A_i$ $(F_{(t)} \leqq (s+1,\dots,m))$, $G_{(t+1)} \leqq (s+1,\dots,n)$. Add $\sum_{i=1}^{s} A_i$ to both sides; then, owing to (3), $\sum_{i=1}^{s} A_i + \sum_{j\in G_{(t+1)}} B_j \leqq \sum_{i\in F_{(s+t)}} A_i$, obtains, where $F'_{(s+t)} = (1,\dots,s) + F_{(t)}$. Now (1) gives $\sum_{j\in G_{(s+t+1)}} B_j \leqq \sum_{i\in F'_{(s+t)}} A_i$, where $G'_{(s+t+1)} = (1,\dots,s) + G_{(t+1)}$, contradicting (b) for the original system.

Thus all possibilities are exhausted. Now the $m+n$ of (4) is $2s$ and that of (5) is $m+n-2s$. As $s \leqq \mathrm{Min}(m,n)$ but not $s=m=n$, so $2s < m+n$, as $s > 0$ so $m+n-2s < m+n$. Hence (owing to our hypothesis concerning the minimum character of $m+n$) our lemma is true both for (4) and for (5).

So we have two permutations $(1',\dots,s')$ and $(1'',\dots,s'')$ of $(1,\dots,s)$ as well as a permutation $((s+1)',\dots,m')$ of $(s+1,\dots,m)$ and a permutation $((s+1)'',\dots,n'')$ of $(s+1,\dots,n)$ together with a $\overline{p} = 0,1,\dots,s$, and a $\overline{q} = 0,1,\dots,\mathrm{Min}(m,n)-s$, such that the equivalents of (α)–(γ) hold for both (4) and (5). That is:

$$(6) \qquad A'_{k'} \cdot B_{k''} \neq \theta \quad \text{for } k = 1,\dots,\overline{p},$$

$$(7) \qquad A_{k'} \cdot B'_{k''} \neq \theta \quad \text{for } k = s+1,\dots,s+\overline{q},$$

$$(8) \qquad \text{whenever } A'_i \leqq B' \ (i=1,\dots,s), \quad \text{then } i = k' \text{ for } k = 1,\dots,\overline{p},$$

$$(9) \qquad \text{whenever } A_i \leqq B^2 \ (i=s+1,\dots,m), \quad \text{then } i = k' \text{ for } k = s+1\dots,s+\overline{q},$$

$$(10) \qquad \text{whenever } B_j \leqq A^1 \ (j=1,\dots,s), \quad \text{then } j = k'' \text{ for } k = 1,\dots,\overline{p},$$

(11)

$$\text{whenever } B'_j \leqq A^2 \ (j=s+1,\dots,n), \quad \text{then } j = k'' \text{ for } k = s+1,\dots,s+\overline{q}.$$

Observe first that (6), (7) imply

$$(12) \qquad A_{k'} \cdot B_{k''} \neq \theta \quad \text{for } k = 1,\dots,\overline{p} \text{ or } s+1,\dots,s+\overline{q}.$$

We shall now consider the consequences of (8)–(11). Consider these two hypotheses:

(13) $$A_i \leqq B \qquad (i = 1, \ldots, m),$$

(14) $$B_j \leqq A \qquad (j = 1, \ldots, n).$$

We discuss all possibilities:

(13) and $i = 1, \ldots, s$. In this case (2) gives

$$A_i' \leqq B\left(-\sum_{j=s+1}^{n} B_j\right) = \left(\sum_{j=1}^{n} B_j\right)\left(-\sum_{j=s+1}^{n} B_j\right) = \sum_{j=1}^{s} B_j = B^1$$

(use (4)); so by (8) $i = k'$ for $k = 1, \ldots, \overline{p}$.

(13) and $i = s+1, \ldots, m$. In this case $A_i \leqq \sum_{j=s+1}^{m} A_i \leqq \sum_{i=1}^{s} A_i$; hence

$$A_i \leqq B\left(-\sum_{i=1}^{s} A_i\right) = \left(\sum_{j=1}^{n} B_j\right)\left(-\sum_{i=1}^{s} A_i\right)$$
$$= \left(\sum_{j=1}^{s} B_j\right)\left(-\sum_{i=1}^{s} A_i\right) + \left(\sum_{j=s+1}^{n} B_j\right)\left(-\sum_{i=1}^{s} A_i\right).$$

Now the first term is $= \theta$ by (1), and the second term is $\sum_{j=s+1}^{m} B_j'$ by (3). So we have $A_i \leqq B^2$ (use (5)); so by (9) $i = k'$ for $k = s+1, \ldots, s+\overline{q}$.

(14) and $j = 1, \ldots, s$. Owing to (1) we have (for $j = 1, \ldots, s$, even without assuming (14))

$$B_j \leqq \sum_{j=1}^{s} B_j \leqq \sum_{i=1}^{s} A_i.$$

Also

$$B_j \leqq \sum_{j=1}^{s} B_j \leqq -\sum_{j=s+1}^{n} B_j.$$

So

$$B_j \leqq \left(\sum_{i=1}^{s} A_i\right)\left(-\sum_{j=s+1}^{n} B_j\right).$$

This is $= \sum_{i=1}^{s} A_i'$, by (2). So we have $B_j \leqq A^1$ (use (4)); so by (10) $j = k''$, for $k = 1, \ldots, \overline{p}$.

We may add this observation: Since, as observed above, this is true for every $j = 1, \ldots, s$, the number of the k'' with $k = 1, \ldots, \overline{p}$ is s. So we have

(15) $$\overline{p} = s.$$

(14) and $j = s+1, \ldots, n$. In this case (3) gives

$$B_j' \leqq A\left(-\sum_{i=1}^{s} A_i\right) = \left(\sum_{i=1}^{m} A_i\right)\left(-\sum_{i=1}^{s} A_i\right) = \sum_{i=s+1}^{m} A_i = A^2$$

(use (5)); so by (ii) $j = k''$ for $k = s+1, \ldots, s+\overline{q}$.

Summing up: Owing to (15), (12) now states

(16) $$A_{k'} \cdot B_{k''} \neq \theta \quad \text{for } k = 1, \ldots, s + \overline{q},$$

and the above four cases give

(17) $$\text{(13) implies } i = k' \quad \text{for } k = 1, \ldots, s + \overline{q},$$

(18) $$\text{(14) implies } j = k'' \quad \text{for } k = 1, \ldots, s + \overline{q}.$$

Thus (α)–(γ) are satisfied (by (16)–(18)) with $p = s + \overline{q}$. This contradicts our original assumption.

For (B): $(*)$ is always true. Consider an $A_i \cdot B_j \neq \theta$ following (II), say

(19) $$A_i B_1 \neq \theta.$$

Put

(20) $$A_i^* = A_i(-B_1) \quad \text{for } i = 2, \ldots, m,$$

(21) $$B_j^* = B_j(-A_1) \quad \text{for } j = 2, \ldots, n.$$

We form the system

(22) $$A^* = \sum_{i=2}^{m} A_i^*, \qquad B^* = \sum_{j=2}^{n} B_j^*.$$

We claim that it satisfies the hypotheses (b), (c) of our lemma. By symmetry, it suffices to consider (b). Assume therefore that (b) is violated:

$$\sum_{i \in F_{(s)}} A_i^* \leqq \sum_{j \in G_{(s+1)}} B_j^* \qquad (F_{(s)} \leqq (2, \ldots, m), G_{(s+1)} \leqq (2, \ldots, n)).$$

Then, *a fortiori* $\sum_{i \in F_{(s)}} A_i \geqq \sum_{j \in G_{(s+1)}} B_j^*$. Add A_1 to both sides; then, owing to (21), $\sum_{i \in F'_{(s+1)}} A_i \geqq \sum_{j \in G_{(s+1)}} B_j$, where $F'_{(s+1)} = (1) + F_{(s)}$. Since $(*)$ is always true, this necessitates $s + 1 = m = n$. But $G_{(s+1)} \leqq (2, \ldots, n)$ implies $s + 1 \leqq n - 1$ and thus a contradiction.

Now the $m + n$ of (22) is $m + n - 2$, hence $< m + n$. Hence (owing to our hypothesis concerning the minimum character of $m+n$), our lemma is true for (22).

So we have a permutation $(2', \ldots, m')$ of $(2, \ldots, m)$ and a permutation $(2'', \ldots, n'')$ of $(2, \ldots, n)$ together with a $\tilde{p} = 0, 1, \ldots, \operatorname{Min}(m, n) - 1$, such that the equivalents of (α)–(γ) hold for (22). That is:

(23) $$A_{k'}^* \cdot B_{k''}^* \neq \theta \quad \text{for } k = 2, \ldots, \tilde{p} + 1,$$

(24) $$\text{whenever } A_i^* \leqq B^* \ (i = 2, \ldots, n) \quad \text{then } i = k' \text{ for } k = 2, \ldots, \tilde{p} + 1,$$

(25) $$\text{whenever } B_j^* \leqq A^* \ (j = 2, \ldots, m) \quad \text{then } j = k'' \text{ for } k = 2, \ldots, \tilde{p} + 1.$$

Put $1' = 1'' = 1$. Then $(1', 2', \ldots, m')$ is a permutation of $(1, 2, \ldots, m)$ and $(1'', 2'', \ldots, n'')$ is a permutation of $(1, 2, \ldots, n)$. (19), (23) imply

(26) $$A_{k'} \cdot B_{k''} = \theta \quad \text{for } k = 1, 2, \ldots, \tilde{p} + 1.$$

We shall now consider the consequences of (24), (25). Consider the two hypotheses:

(27) $$A_i \leqq B \qquad (i = 1, \ldots, m),$$

(28) $$B_j \leqq A \qquad (j = 1, \ldots, n).$$

Let us analyze (27). If $i = 1$ then $i = 1'$. If $i \neq 1$ that is $i = 2, \dots, m$, then $A_i \leqq \sum_{i=2}^m A_i \leqq -A_1$, so, *a fortiori*, $A_i^* \leqq -A_i$. On the other hand by (20)

$$A_i^* = A_i(-B_1) \leqq B(-B_1) = \left(\sum_{j=1}^n B_j\right)(-B_1) = \sum_{j=2}^n B_j.$$

Consequently $A_i^* \leqq (\sum_{j=2}^n B_j)(-A_1)$. This is $\sum_{j=2}^n B_j^*$ by (21). So we have $A_i^* \leqq B^*$ (use (22)); so by (24) $i = k'$ for $k = 2, \dots, \tilde{p}+1$.

Summing up,

(29) $\qquad$ (27) implies $i = k'$ for $k = 1, 2, \dots, \tilde{p}+1$.

And now by symmetry

(30) $\qquad$ (28) implies $j = k''$ for $k = 1, 2, \dots, \tilde{p}+1$.

Thus (α)–(γ) are satisfied (by (26), (29), (30)) with $p = \tilde{p}+1$. This contradicts our original assumption.

Thus all alternatives and readers are exhausted, and the proof is complete.

3. Notations (topology and group theory)

G : Topological group.
xy : Composition rule (in G).
x^{-1} : Reciprocal (in G).
1 : Unit (in G).
M, N : Arbitrary subset of G.
O, P, Q : : Open subset of G.
C, D, E : Compact subsets of G.
$\overline{M}$: Closure of M (in G).
M^i : Interior of M (in G).
xM : Set $(xu \mid u \in M)$.
Mx : Set $(ux \mid u \in M)$.
M^{-1} : Set $(u^{-1} \mid u \in M)$.

Hypotheses:
(1) x^{-1} is a continuous (1-variable) function of x (in all G).
(2) xy is a continuous (2-variable) function of x, y (in all G).
(3) G is locally compact; i.e., there exists a C with $1 \in C^i$.

4. Equidistribution

Let a C be given which will remain fixed throughout all our discussions.
We define:

Definition. A set

(1) $$\overline{F} = \overline{F}_{(m)} = (a_1, \dots, a_m) \leqq C \qquad (m = 1, 2, \dots)$$

is O-*equidistributed* ($1 \in O$) if and only if it possesses this property: For every $a \in G$ there exists a permutation $(\bar{1}, \dots, \bar{m})$ of $(1, \dots, m)$, such that for every $i = 1, \dots, m$ either

(A) $a_i O \cdot a a_{\bar{i}} O \neq \theta$ or

(B) $a_i O \cdot (-aC) \neq \theta$ and $a_{\bar{i}} O(-a^{-i} C) \neq \theta$.

We prove, for future application (in (15)):

Lemma. *Let a P with $P \leqq C$, $P \neq \theta$ be given. Then there exist an $O_0 = O_0(P)$ with ($1 \in O_0$) and a $\beta_0 = \beta_0(C, P) > 0$, with the following properties:*

If the $\overline{F}$ of (1) above is O-equidistributed, with an O with $1 \leqq O \leqq O_0$, then the number of i $(= 1, \dots, m)$ with $a_i \in P$ (i.e. the number of elements of $\overline{F} \cdot P$) is $\geqq \beta_0 m$.

Proof. Since xyz^{-1} is continuous (cf. (3)), and since $1 \in P$, there exists an $O_0 = O_0(P)$ with $1 \in O_0$ and $O_0 O_0 O_0^{-1} \leqq P$. Consequently we have

$$1 \in O_0 \leqq O_0 O_0 \leqq O_0 O_0 O_0^{-1} \leqq P \leqq C. \tag{2}$$

Clearly $C \leqq \sum_{d \in C} dO_0 \cdot C$ is compact, every dO_0 open; consequently

$$C \leqq \sum_{n=1}^{k} d_n O_0. \tag{3}$$

k and the $d_1, \dots, d_k$ depend on C, O_0 i.e. on C, P—we note this by writing $k = k(C, P)$.

Now consider an O with $1 \in O \leqq O_0$ and set $\overline{F}$ of the form (1), which is O-equidistributed.

All $a_i \in C$ so (3) implies

$$\sum_{n=1}^{k} (\text{Number of } i \ (= 1, \dots, k) \text{ with } a_i \in d_n O_0) \geqq m.$$

Hence there exists an n $(= 1, \dots, k)$ with

$$(\text{Number of } i \ (= 1, \dots, k) \text{ with } a_i \in d_n O_0) \geqq \frac{m}{k}. \tag{4}$$

Consider an i with $a_i \in d_n O_0$. We must have one of the alternatives (A) and (B) given in the definition above.

(B) would imply $a_i O(-d_n C) \neq \theta$. Since $a_i O \leqq d_n O_0 O \leqq d_n O_0 O_0 \leqq d_n C$ (use (2)), this is impossible.

Hence we have (A). This implies $a_i O \cdot d_n a_{\bar{i}} O \neq \theta$; consequently

$$a_{\bar{i}} \in d_n^{-1} a_i O O^{-1} \leqq d_n^{-1} (d_n O_0) O O^{-1} = O_0 O O^{-1} \leqq O_0 O_0 O_0^{-1} \leqq P$$

(use (2)). So we see:

$$a_i \in d_n O \quad \text{implies} \quad a_{\bar{i}} \in P. \tag{5}$$

Consequently

$$\begin{aligned} &(\text{Number of } j \ (= 1, \dots, k) \text{ with } a_j \in P) \\ &\qquad \geqq (\text{Number of } i \ (= 1, \dots, k) \text{ with } d_i \in d_n \theta). \end{aligned} \tag{6}$$

Combination of (4), (6) gives

$$\frac{(\text{Number of } j \ (= 1, \dots, k) \text{ with } a_j \in P)}{m} \geqq \frac{1}{k} = \frac{1}{k(C, P)}. \tag{7}$$

Thus we can put

$$\beta_0 = \beta_0(C, P) = 1/k = \frac{1}{k(C, P)} > 0$$

and the proof is completed.

5. First example of equidistribution

Consider the C, O of §4; we assume only $1 \in O$. Clearly $C \leqq \sum_{a \in C} aO \cdot C$ is compact; every aO is open; consequently for a suitable finite set

(1) $$\overline{F} = \overline{F}_{(m)} = (a_1, \dots, a_m) \leqq C \qquad (m = 1, 2, \dots)$$

we have

(2) $$C \leqq \sum_{i=1}^{m}{}^{\bullet}\, a_i O.$$

Obviously (2) is equivalent to

(3) $$C = \sum_{i=1}^{m} M_i \quad \text{with } M_i \leqq a_i O.$$

Now we claim:

Lemma. *Choose the $\overline{F}$ of* (1) *with a minimum m satisfying* (2) (*that is* (3)). *This $\overline{F}$ is O-equidistributed.*

PROOF. Use the form (3). We have (for any a)

(4) $$C = \sum_{i=1}^{m} M_i, \qquad aC = \sum_{i=1}^{m} aM_i.$$

Apply the lemma of §2 (of Maak and Kakutani) to (4) in place of its (a). (I.e.: $m = n$, $A = C$, $B = aC$, $A_i = M_i$, $B_i = aM_i$.) We prove the equivalents of its (b), (c), that is

(5) $$\text{never } \sum_{i \in F_{(s)}} M_i \geqq \sum_{i \in G_{(s+1)}} aM_i$$

and

(6) $$\text{never } \sum_{i \in G_{(s)}} aM_i \geqq \sum_{i \in F_{(s+1)}} M_i.$$

Since (5) arises from (6) by multiplying both sides by a^{-1} and then replacing a by a^{-1} it suffices to consider (6). Assume the opposite: (6) is violated. Then

$$\sum_{i \in G_{(s)}} aM_i \geqq \sum_{i \in F_{(s+1)}} M_i;$$

hence (4) gives

$$C \leqq \sum_{i \in (1, \dots, m) - F_{(s+1)}} + {}^{\bullet}\sum_{i \in G_{(s)}} aM_i,$$

so by (3)

$$C \leqq \sum_{i \in (1, \dots, m) - F_{(s+1)}}^{\bullet} a_i O + {}^{\bullet}\sum_{i \in G_{(s)}}^{\bullet} aa_i O.$$

Denote the a_i with $i \in (1, \dots, m) - F_{(s+1)}$ and the aa_i with $i \in G_{(s)}$ in some order, by $\overline{a}_1, \dots, \overline{a}_{m-1}$. Then we have $C \leqq \sum_{k=1}^{m-1} \overline{a}_k O$, contradicting the minimum property of m in the form (2).

Hence the lemma of §2 is true for (4).

So we have two permutations $(1', \dots, m')$ and $(1'', \dots, m'')$ of $(1, \dots, m)$ together with a $p = 0, 1, \dots, m$ such that the equivalents of (α)–(γ) of §2 hold for (4). That is

$$M_{k'} \cdot aM_{k''} \neq \theta \quad \text{for } k = 1, \dots, p, \tag{7}$$

$$\text{whenever } M_i \leqq aC, \quad \text{then } i = k' \text{ for } k = 1, \dots, p, \tag{8}$$

$$\text{whenever } aM_i \leqq C, \quad \text{then } i = k'' \text{ for } k = 1, \dots, p. \tag{9}$$

Combining (3) and (7)

$$a_{k'} \cdot O \cdot aa_{k''} O \neq \theta \quad \text{for } k = 1, \dots, p \tag{10}$$

obtains.

Next $a_{\overline{k}} O \leqq aC$ or $a_{k''} O \leqq a^{-1} C$ $(k = 1, \dots, m)$ imply by (3) that $M_{k'} \leqq aC$ or $aM_{k''} \leqq C$; hence by (8) and (9) at any rate $k = 1, \dots, p$. So we see

$$a_{k'}(O)(-aC) \neq \theta \quad \text{and} \quad a_{k''} O(-a^{-1}C) \neq \theta \tag{11}$$

for $k = p + 1, \dots, m$.

Thus the alternative of (A) or (B) in the definition of §4 is guaranteed (by (10), (11)), if we define the permutation $(\overline{1}, \dots, \overline{m})$ of §4 by $\overline{k}' = k'$ for $k = 1, \dots, m$.

6. Second example of equidistribution

Consider the C, O of §4. We now assume $1 \in O$ and $bO \leqq C$ for some b.

Consider those finite sets

$$\overline{G} = \overline{G}_{(n)} = (b_1, \dots, b_n) \leqq C \qquad (n = 1, 2, \dots) \tag{1}$$

for which we have

$$C \geqq \sum_{j=1}^{n} b_j O. \tag{2}$$

(The main point is the disjointness of the addends $b_j O$.) Obviously (2) is equivalent to

$$C = \sum_{j=1}^{n} M_j \quad \text{with } M_j \geqq a_j O. \tag{3}$$

Such sets $\overline{G}$ exist by virtue of our original assumptions: e.g. $\overline{G} = \overline{G}_{(1)} = (b)$ (cf. above). Now we claim

Lemma. *Choose the $\overline{G}$ of (1) with a maximum n satisfying (2) (that is (3)). This is possible (that is, the n in question are bounded). This $\overline{G}$ is $OO^{-1}O$-equidistributed.*

PROOF. The n in question are bounded. Since $x^{-1}y$ is continuous (cf. §3), we can choose a P with $1 \in P$ and $PP^{-1} \leqq O$. Apply the considerations at the beginning of §5 to P in place of O; then we obtain

$$C \leqq \sum_{i=1}^{m}{}^{\bullet}\, a_i P, \tag{4}$$

with a certain m which we consider as fixed.

Consider now (2), assuming $n > m$. Then, as $1 \in O$ all $b_j \in C$ for $j = 1, \ldots, n$. Hence there must exist two different $j_1, j_2 = 1, \ldots, n$ with $b_{j_1}, b_{j_2} \in a_i P$ for the same $i = 1, \ldots, m$. Consequently $a_i \in b_{j_1} P^{-1}$; hence $b_{j_2} \in a_i P \leqq b_{j_1} P^{-1} P \leqq b_{j_1} O$. Since $b_{j_2} \in b_{j_2} O$, this contradicts $b_{j_1} O \cdot b_{j_2} O = \theta$, that is (2).

Thus necessarily $n \leqq m$; that is our n are bounded.

$\overline{G}$ is $OO^{-1}O$-equidistributed: Use the form (2). Consider a b for which neither

$$bO(-C) \neq \theta \tag{5}$$

nor

$$b \in b_j OO^{-1} \quad \text{for some } j = 1, \ldots, n. \tag{6}$$

For this b then $bO \leqq C$ and $b \notin b_j OO^{-1}$, $bO \cdot b_j O = \theta$ for all $j = 1, \ldots, n$. Hence

$$C \geqq \sum_{j=1}^{n} b_j O + bO, \tag{7}$$

and so (7) would give (2) with $n + 1$ (put $b_{(n+1)} = b$) contradicting the maximum property of n. Thus every b fulfills (5) or (6); that is, if we define

$$C' = \sum_{j=1}^{n}{}^{\bullet}\, b_j OO^{-1}, \tag{8}$$

then

$$b \notin C' \quad \text{implies} \quad bO(-C) \neq \theta. \tag{9}$$

Since $b \notin C$ trivially implies $bO(-C) \neq \theta$, we can strengthen (9) to

$$b \notin C \cdot C^1 \quad \text{implies} \quad bO(-C) \neq \theta. \tag{10}$$

Now we have, owing to (2), (8) and to the obvious fact $b_j O \leqq b_j OO^{-1}$ (remember $1 \in O$), these relations

$$\sum_{j=1}^{n} b_j O \leqq \sum_{j=1}^{n} b_j OO^{-1} \begin{cases} = C', \\ \leqq C; \end{cases}$$

hence

$$\sum_{j=1}^{n} b_j O \leqq C \cdot C' \leqq \sum_{j=1}^{n}{}^{\bullet}\, b_j OO^{-1}. \tag{11}$$

Consequently we can put

$$C \cdot C' = \sum_{j=1}^{n} N_j \quad \text{with } b_j O \leqq N_j \leqq b_j OO^{-1}. \tag{12}$$

We have (for any a)

$$CC' = \sum_{j=1}^{n} N_j, \qquad a(CC') = \sum_{j=1}^{n} aN_j. \tag{13}$$

Apply the lemma of §2 (of Maak and Kakutani) to (13) in place of its (a). (I.e.: $m = n$, $A = CC', B = a(C \cdot C'), A_j = N_j$, $B_j = aN_j$). We prove the equivalents of its (b), (c); that is

$$(13\tfrac{1}{3}) \qquad \text{never } \sum_{j \in F_{(s)}} N_j \geqq \sum_{j \in G_{(s+1)}} aN_j,$$

and

$$(13\tfrac{2}{3}) \qquad \text{never } \sum_{j \in G_{(s)}} aN_j \geqq \sum_{j \in F_{(s+1)}} N_j.$$

Since $(13\frac{2}{3})$ arises from $(13\frac{1}{3})$ by multiplying both sides by a^{-1} and then replacing a by a^{-1} it suffices to consider $(13\frac{1}{3})$. Assume the opposite: $(13\frac{1}{3})$ is violated: Then $\sum_{j \in F_{(s)}} N_j \geqq \sum_{j \in G_{(s+1)}} aN_j$; hence (13) gives

$$\begin{aligned} C \geqq C \cdot C' = \sum_{j=1}^{n} N_j &= \sum_{j \in (1,\ldots,n) - F_{(s)}} N_j + \sum_{j \in F_{(s)}} N_j \\ &\geqq \sum_{j \in (1,\ldots,n) - F_{(s)}} N_j + \sum_{j \in G_{(s+1)}} aN_j, \end{aligned}$$

so by (12)

$$C \geqq \sum_{j \in (1,\ldots,n) - F_{(s)}} b_j O + \sum_{j \in G_{(s+1)}} ab_j O. \tag{14}$$

Denote the b_j with $j \in (1, \ldots, n) - F_{(s)}$ and the ab_j with $j \in G_{(s+1)}$ in some order, by $\bar{b}_1, \ldots, \bar{b}_{n+1}$. Then

$$(14\tfrac{1}{2}) \qquad C \geqq \sum_{k=1}^{n+1} \bar{b}_k O,$$

contradicting the maximum property of n in the form (2).

Hence the lemma of §2 is true for (13).

So we have two permutations $(1', \ldots, n')$ and $(1'', \ldots, n'')$ together with a $p = 0, 1, \ldots, n$ such that the equivalents of (α)–(γ) of §2 hold for (13). That is:

$$N_{k'} \cdot aN_{k''} \neq \theta \quad \text{for } k = 1, \ldots, p, \tag{15}$$

$$\text{whenever } N_j \leqq a(C \cdot C'), \quad \text{then } j = k' \text{ for } k = 1, \ldots, p, \tag{16}$$

$$\text{whenever } aN_j \leqq C \cdot C', \quad \text{then } j = k'' \text{ for } k = 1, \ldots, p. \tag{16$'$}$$

Combining (12) and (15) we obtain

$$a_{k'} OO^{-1} \cdot a_{k''} OO^{-1} \neq \theta \quad \text{for } k = 1, \ldots, p. \tag{17}$$

Next, $a_{k'}OO^{-1} \leqq a(C \cdot C')$ or $a_{k''}OO^{-1} \leqq a^{-1}(C \cdot C')$ implies by (12) that $N_{k'} \leqq a(C \cdot C')$ or $aN_{k''} \leqq C \cdot C'$; hence by (16) and (16′) at any rate $k = 1, \ldots, p$. On the other hand $a_{k'}OO^{-1} \leqq a(C \cdot C')$ or $a_{k''}OO^{-1} \leqq a^{-1}(C \cdot C')$ means $a^{-1}a_{k'}OO^{-1} \leqq C$

or $aa_{k''}OO^{-1} \leqq C \cdot C'$; hence by (10) they are implied by $a^{-1}a_{k'}OO^{-1}O \leqq C$ or $aa_{k''}OO^{-1}O \leqq C$, that is by $a_{k'}OO^{-1}O \leqq aC$ or $a_{k''}OO^{-1}O \leqq a^{-1}C$. So we see

$$(18) \qquad a_{k'}OO^{-1}O \cdot (-aC) \neq \theta \quad \text{and} \quad a_{k''}OO^{-1}O(-a^{-1}C) \neq \theta$$

for $k = p+1, \dots, n$.

Thus the alternative (A) or (B) in the definition of §4 is guaranteed for $OO^{-1}O$ in place of its O (by (17), (18), remember $OO^{-1} \leqq OO^{-1}O$ since $1 \in O$), if we define the permutation $(\overline{1}, \dots, \overline{n})$ by $\overline{k'} = k''$ for $k = 1, \dots, n$.

7. Equidistribution (concluded)

Consider the C, O of §4. We assume only $1 \in O$. Then the lemma of §5 gives immediately an O-equidistributed set

$$\overline{F} = \overline{F}_{(m)} = (a_1, \dots, a_m).$$

In order to apply the lemma of §6, we must assume $C^i \neq \theta$. Choose a $b \in C^i$; then $1 \in P = b^{-1}C^i$, and $bP \leqq C$. Since $xy^{-1}z$ is continuous (see §3), we can choose a Q with $1 \in Q$ and $QQ^{-1}Q \leqq O$. Put $O' = PQ$. Then

$$(1) \qquad 1 \in O', \quad bO' \leqq C, \quad O'O'^{-1}O' \leqq O.$$

So we can apply the lemma of §6 to O' in place of its O and it gives an O-equidistributed set $\overline{G} = \overline{G}_{(n)} = (b_1, \dots, b_n)$.

8. Continuous functions

We define

Definition. Let a set M ($\leqq G$) be given. $\mathcal{F}_M$ is the system of those functions $f(x)$ which possess the following properties:

(α) $f(x)$ is defined for all x (in G).
(β) The values of $f(x)$ are real numbers.
(γ) $f(x)$ is a continuous (1-variable) function of x (in all G).
(δ) $f(x) = 0$ whenever $x \notin M$.

$\overline{\mathcal{F}}$ is the system of those functions $f(x)$ which possess the above properties (α)–(δ) except that (β) is replaced by the (stronger) requirement

($\overline{\beta}$) The values of $f(x)$ are real numbers $\geqq 0$, $\leqq 1$.

Since $f(x)$ is continuous, $f(x) = 0$ in all $-M$ implies the same in all $-\overline{M}$; i.e. (δ) above extends from M to $-\overline{(-M)} = M^i$. So we see

$$(1) \qquad \mathcal{F}_M = \mathcal{F}_{M^i}, \qquad \overline{\mathcal{F}}_M = \overline{\mathcal{F}}_{M^i}.$$

For any function $f(x)$ we define

$$(2) \qquad f^{\checkmark}(x) = f(x^{-1}).$$

Clearly $f \in \mathcal{F}_M$ ($\overline{\mathcal{F}}_M$) is equivalent to $f^{\checkmark} \in \mathcal{F}_{M^{-1}}$ ($\overline{\mathcal{F}}_{M^{-1}}$). We also define

$$(3) \qquad \|f\| = \sup_{x \in G} |f(x)|,$$

$$(4) \qquad \mathrm{Osc}_O(f) = \sum_{x^{-1}y \in O} |f(x) - f(y)| \qquad (1 \in O).$$

It is well known that both these quantities are finite. Clearly

$$(5) \qquad \|f^{\checkmark}\| = \|f\|.$$

Combination of (2) and (4) gives

(6) $$\operatorname{Osc}_O(f^{\checkmark}) = \sup_{xy^{-1}\in O} |f(x) - f(y)|.$$

(Compare (4) and (6)!) Interchanging x, y in (4) gives, since $y^{-1}x = (x^{-1}y)^{-1}$,

(7) $$\operatorname{Osc}_O(f) = \operatorname{Osc}_{O^{-1}}(f).$$

(3) and (4) give at once

(8) $$\operatorname{Osc}_O(f) = \operatorname{Osc}_{O^{-1}}(f) \leqq 2\|f\|.$$

We also state some well-known computation rules, involving $\|\cdots\|$ and Osc: (c a real constant)

(9) $$\begin{cases} \|f+g\| \leqq \|f\| + \|g\|, & \|cf\| = |c|\cdot\|f\|, \\ \|fg\| \leqq \|f\|\cdot\|g\|, \end{cases}$$

(10) $$\begin{cases} \operatorname{Osc}_O(f+g) \leqq \operatorname{Osc}_O(f) + \operatorname{Osc}_O(g), & \operatorname{Osc}_O(cf) = |c| \operatorname{Osc}_O(f), \\ \operatorname{Osc}_O(fg) \leqq \|f\| \operatorname{Osc}_O(g) + \|g\| \operatorname{Osc}_O(f). \end{cases}$$

We prove now some lemmas concerning $\|\cdots\|$ and Osc in connection with the C of §4. These deal with well-known properties of continuity and compactness, and the proofs are only given for the sake of completeness.

LEMMA I. *For an $f \in \mathcal{F}_C$ and any $\varepsilon > 0$ there exists an $O = O(f, \varepsilon)$ with $1 \in O$ and* $\operatorname{Osc}_O(f) \leqq \varepsilon$.

PROOF. For every $a \in C$ there exists an O_a with $a \in O$, such that $x \in O_a$ implies $|f(x) - f(a)| \leqq \frac{1}{2}\varepsilon$.

Since auv is continuous (cf. §3, with respect to u, v, we consider a as fixed for the moment), we can choose a P_a with $1 \in P_a$ and a $P_a P_a \leqq O_a$.

Now $C \leqq \sum_{a\in C} aPa$. Since C is compact and every aP_a open, $C \leqq \sum_{i=1}^k a_i P_{a_i}$. Put $O = \prod_{i=1}^k P_{a_i} \cdot P_{a_i}^{-1}$. Then clearly $1 \in O$ and $O = O^{-1}$.

Assume first $x^{-1}y \in O$ and $x \in C$. Then $x \in a_i P_{a_i}$ for some $i = 1, \ldots, k$. Hence, *a fortiori*, $x \in a_i P_{a_i} \cdot P_{a_i} \leqq O_{a_i}$. Next $y = xx^{-1}y \in a_i P_{a_i} O \leqq a_i P_{a_i} \cdot P_{a_i} \leqq O_{a_i}$. So $|f(x) - f(a_i)|$, $|f(y) - f(a_i)| \leqq \frac{1}{2}\varepsilon$. Consequently

(11) $$|f(x) - f(y)| \leqq \varepsilon.$$

Assume next $x^{-1}y \in O$ and $y \in C$. Since $y^{-1}x = (x^{-1}y)^{-1} \in O^{-1} = O$, we obtain again (11), by interchanging x, y. Assume finally $x^{-1}y \in O$, and $x, y \notin C$. Then $f(x) = f(y) = 0$; hence (11) is still true.

So (11) whenever $x^{-1}y \in O$. Consequently (cf. (4)) $\operatorname{Osc}_O(f) \leqq \varepsilon$. Thus the proof is completed.

LEMMA II. *For an $f \in \mathcal{F}_C$ the $\|f\|$ of* (3) *and the* $\operatorname{Osc}_O(f)$ *of* (4) *are finite.*

PROOF. Owing to (7) it suffices to consider $\|f\|$. Form the O of Lemma I above for $\varepsilon = 1$. Since $C \leqq \sum_{a\in C} aO$, we have again (cf. above) $C \leqq \sum_{i=1}^k a_i O$. Consider an $x \in C$. Then $x \in a_i O$ for some $i = 1, \ldots, k$. Hence $a_i^{-1}x \in O$, $|f(a_i) - f(x)| \leqq 1$, $|f(x)| \leqq |f(a_i)| + 1$. Consequently $x \in C$ implies

(12) $$|f(x)| \leqq \alpha \text{ for the constant } \alpha = \operatorname{Max}_{i=1,\ldots,k} |f(a_i)| + 1.$$

For $x \notin C$ we have $f(x) = 0$; hence (12) is still true. So (12) for all x and therefore $\|f\|$ is finite.

LEMMA III. *For an $f \in \mathcal{F}_C$ the $\|f\|$ of (3) is not merely an l.u.b., but it is an assumed maximum.*

PROOF. Assume the opposite; then $|f(x)| < \|f\|$ for all x. Consequently

$$g(x) = \frac{1}{\|f(x)\| - f(x)}$$

is also a function in $\mathcal{F}_C$. Hence Lemma II permits us to form the finite $\|g\|$. Now $|g(x)| \leqq \|g\|$, i.e. $1/|\,\|f(x)\| - f(x)| \leqq \|g\|$, $f(x) \leqq \|f\| - 1/\|g\|$ for all x. Consequently $\|f\| \leqq \|f\| - 1/\|g\|$ and this is obviously impossible.

LEMMA IV. *For every O with $1 \in O$ there exists an $O^{\circledast} = O^{\circledast}(C, O)$ (independent of f cf. below), with $1 \in O^{\circledast} \leqq O$ which possesses the following property:*

$$\textit{For every } f \in \mathcal{F}_C, \qquad \mathrm{Osc}_{O^{\circledast}}(f^{\checkmark}) \leqq \mathrm{Osc}_O(f).$$

PROOF. Comparing (4) and (6), and remembering that $f(z) = 0$ for $z \notin C$—i.e. $f^{\checkmark}(z) = f(z^{-1}) = 0$ for $z \notin C^{-1}$—we see that we need an $O^{\circledast}$ with $1 \in O^{\circledast} \leqq O$ and the following property:

$$(13) \qquad yx^{-1} \in O^{\circledast} \text{ and } x \in C^{-1} \text{ or } y \in C^{-1} \quad \text{imply} \quad x^{-1}y \in O.$$

Observe that

$$(14) \qquad x^{-1}y = x^{-1} \cdot yx^{-1} \cdot x = y^{-1} \cdot yx^{-1} \cdot x.$$

Since $u^{-1}vu$ is continuous (cf. §3, with respect to u, v), and since it is 1 for $v = 1$ (and any u), there exist for every a, O_a, P_a with $a \in O_a$, $1 \in P_a$ such that $u \in O_a$, $v \in P_a$ imply $u^{-1}vu \in O$. Since $C^{-1} \leqq \sum_{a \in C^{-1}} O_a$, we have again (cf. above) $C^{-1} \leqq \sum_{i=1}^{k} O_{a_i} \ldots$. Put $O^{\circledast} = \prod_{i=1}^{k} P_{a_i} \cdot O$. Then $1 \in O^{\circledast} \leqq O$.

Assume $u \in C^{-1}$, $v \in O^{\circledast}$. Then $u \in O_{a_i}$ for some $i = 1, \ldots, k$ and $v \in P_{a_i}$. Hence $u^{-1}vu \in O$. So we see

$$(15) \qquad u \in C^{-1}, \; v \in O^{\circledast} \quad \text{imply} \quad u^{-1}vu \in O.$$

Combining (15) with (14) gives (13); we must only put $v = yx^{-1}$ and $u = x$ or $u = y$. This completes the proof.

9. Means

Consider the C of §4. We define

Definition. A *mean* is a functional $\mathbb{M}(f) = \mathbb{M}_x(f(x))$, which possesses the following properties:

(α) $\mathbb{M}(f)$ is defined (precisely) for all $f \in \mathcal{F}_C$.
(β) $\mathbb{M}(f + g) = \mathbb{M}(f) + \mathbb{M}(g)$.
(γ) $\mathbb{M}(f) \geqq 0$ for $f \in \overline{\mathcal{F}}_C$.

The three properties (δ)–(η) which we shall state now would have been postulated within Definition I if they were not consequences of (α)–(γ). Since they are, we will prove them.

LEMMA I. *We have for every mean $\mathbb{M}(f)$,*
(δ) *$\mathbb{M}(kf) = k\mathbb{M}(f)$ (k any real constant)*
(ε) *$\mathbb{M}(f) \geqq 0$ if $f(x) \geqq 0$ for all $x \in G$.*
(η) *There exists a constant $\alpha = \alpha(\mathbb{M})$ such that for all $f \in \mathcal{F}_C$*

$$|\mathbb{M}(f)| \leqq \alpha\|f\|.$$

PROOF. Ad (δ) for rational k: If (δ) holds for k', k'' where $k'' \neq 0$, then it holds clearly for k'/k'' too. By (β) it holds for $k' \neq k''$ too. Finally it is obviously true for $k = 1$. From these we can conclude that (δ) holds for all rational k.

Ad (ε): Assume $f(x) \geqq 0$ for all $x \in G$. Choose a rational $k > 0$ with $k \geqq \|f\|$, i.e. $0 \leqq f(x) \leqq k$ for all $x \in G$. So $\frac{1}{k}f \in \overline{\mathcal{F}}_C$; hence $\mathbb{M}(\frac{1}{k}f) \geqq 0$ and so by the above $\mathbb{M}(f) \geqq 0$.

Ad (η): Assume the opposite. Then we could find for every η ($= 1, 2, \ldots$) an f_n ($\in \mathcal{F}_C$) with

(1) $$|\mathbb{M}(f_n)| > 4^n \|f_n\|.$$

(1) excludes $f_n = 0$ (then both sides would vanish); hence $\|f_n\| > 0$. Multiply f_n by $1/2^n\|f_n\|$—then (1) remains true, and $\|f_n\|$ becomes equal to $1/2^n$. I.e., we have with this new choice of f_n

(2) $$|\mathbb{M}(f_n)| > 2^n$$

and $\|f_n\| = 1/2^n$, and so $|f_n(x)| \leqq 1/2^n$ for all $x \in G$. Thus we can form

$$g(x) = \sum_{n=1}^{\infty} |f_n(x)|,$$

and it belongs to $\mathcal{F}_C$. We have $g(x) \geqq |f_n(x)|$; hence $g(x) \geqq f_n(x)$, and so by (β) and (ε) $\mathbb{M}(g) \geqq \mathbb{M}(f_n)$, i.e. by (2) $\mathbb{M}(g) > 2^n$. This must be true for all $n = 1, 2, \ldots$ which is obviously impossible. Consequently (η) is true.

Ad (δ) for all k: For a fixed f and variable k, $\mathbb{M}(kf)$ is a continuous function of k at $k = 0$: (η) gives $|\mathbb{M}(kf)| \leqq \alpha\|kf\| = \alpha\|f\| \cdot |k|$. It is clearly additive in k; hence it is continuous for all k. We have already proved (δ) for the rational k; now it extends by continuity to all k.

The α of (η) above (for a given mean $\mathbb{M}(f)$) are clearly all $\geqq 0$ and they possess a minimum:

(3) Denote the smallest α of (η) above by $|||\mathbb{M}|||$.

Observe:

LEMMA II. *We have for every mean* $\mathbb{M}(f)$:
(η') $|||\mathbb{M}||| = \sup_{\|f\| \leqq 1} |\mathbb{M}(f)| = \sup_{\|f\|=1} |\mathbb{M}(f)|$.
(η'') $|||\mathbb{M}||| = \sup_{f \in \overline{\mathcal{F}}_C} \mathbb{M}(f)$.

PROOF. Ad (η') : $|||\mathbb{M}|||$ is the smallest α which fulfills (η) above. Since (η) is identically true for $f = 0$ we may restrict it to $\|f\| > 0$. Since it is unaffected by replacement of f by kf (cf. (δ) above), we may even assume $\|f\| = 1$. So $|||\mathbb{M}|||$ is the smallest α with $|\mathbb{M}(f)| \leqq \alpha$ for all $\|f\| = 1$; i.e.

(4) $$|||\mathbb{M}||| = \sup_{\|f\|=1} |M(f)|.$$

Obviously

$$\sup_{\|f\| \leqq 1} |\mathbb{M}(f)| \geqq \sup_{\|f\|=1} |\mathbb{M}(f)|;$$

hence by (4)

(5) $$\sup_{\|f\| \leqq 1} |\mathbb{M}(f)| \geqq |||\mathbb{M}|||.$$

The definition of $|||\mathbb{M}|||$ by (η), on the other hand, gives $|\mathbb{M}(f)| \leqq |||\mathbb{M}|||$ whenever $\|f\| \leqq 1$; hence

$$(6) \qquad \sup_{\|f\| \leqq 1} |\mathbb{M}(f)| \leqq |||M|||.$$

Now (4), (5), (6) together prove (η').

Ad (η''): $f \in \overline{\mathcal{F}}_C$ implies $\|f\| \leqq 1$ and $\mathbb{M}(f) \leqq |\mathbb{M}(f)|$ so

$$(7) \qquad \sup_{f \in \overline{\mathcal{F}}_C} \mathbb{M}(f) \leqq \sup_{\|f\| \leqq 1} |\mathbb{M}(f)|.$$

Consider now an $f \in \mathcal{F}_C$ with $\|f\| \leqq 1$. Put $g(x) = |f(x)|$. Then $0 \leqq g(x) \leqq 1$ for all $x \in G$, i.e. $g \in \overline{\mathcal{F}}_C$. Besides $\pm f(x) \leqq g(x)$ for all $x \in G$, so $(\beta), (\delta), (\varepsilon)$ gives $\pm\mathbb{M}(f) \leqq \mathbb{M}(g)$, i.e. $|\mathbb{M}(f)| \leqq \mathbb{M}(g) \leqq \sup_{h \in \overline{\mathcal{F}}_C} \mathbb{M}(h)$. Thus

$$(8) \qquad \sup_{\|f\| \leqq 1} |\mathbb{M}(f)| \leqq \sup_{h \in \overline{\mathcal{F}}_C} \mathbb{M}(h)$$

Now (7), (8) and (η') prove (η'').

LEMMA III. $\mathbb{M}(f)$ *is identically* 0 *if and only if it is* 0 *for all* $f \in \overline{\mathcal{F}}_C$.

PROOF. Immediate by (η'') above.

10. Left invariance of means

Consider the C, O of §4, assuming only $1 \in O$. We now establish an intrinsic connection between means and the group structure of G.

DEFINITION I. A *mean* $\mathbb{M}(f)$ is *left invariant* (abbreviated l.i.) if it possesses the following property:

(ϑ) $\mathbb{M}_x(f(ax)) = M_x(f(x))$ whenever $f(ax), f(x)$ both belong to $\mathcal{F}_C$.

DEFINITION II. A *mean* $\mathbb{M}(f)$ is *O-approximately left invariant* (abbreviated: O-a.l.i.) if it possesses the following property:

(ϑ_O^*) $|\mathbb{M}_x(f(ax)) - \mathbb{M}_x(f(x))| \leqq K_0 \operatorname{Osc}_O(f)$, whenever $f(ax), f(x)$ both belong to $\overline{\mathcal{F}}_C$.

Observe that (O_O^*) is weaker than (ϑ) for two reasons: First, because it replaces equality by the evaluation $|\cdots| \leqq K_0 \operatorname{Osc}_O(f)$. Second, because it replaces $\mathcal{F}_C$ by $\mathcal{F}_{\overline{C}}$. Cf., however, the lemma below.

The real number K_0 in (ϑ_O^*) in the above definition is an absolute constant, which will be chosen in Lemma I in §15 as $K_0 = 2$. We prefer, however, to carry it along as K_0 because its numerical value is unimportant in our other discussions.

The relationship of Definitions I and II is elucidated by this lemma:

Lemma. *A mean* $\mathbb{M}(f)$ *is l.i. if and only if it is O-a.l.i. for every* O *(with* $1 \in O$*).*

PROOF. Necessity: Obvious, since (ϑ_O^*) is a consequence of (ϑ) (cf. the remark after Definition II).

Sufficiency: Let $\mathbb{M}(f)$ be O-a.l.i. for all O (with $1 \in O$). Consider an f such that $f(ax), f(x)$ both belong to $\mathcal{F}_C$. Hence

$$|\mathbb{M}_x(f(ax)) - \mathbb{M}_x(f(x))| \leqq K_0 \operatorname{Osc}_O(f).$$

Now Lemma I in §8 gives

$$(1) \qquad \mathbb{M}_x(f(ax)) = \mathbb{M}_x(f(x)).$$

Assume next merely $f(x) \geqq 0$ for all $x \in G$. Choose $K > 0$ with $K \geqq \|f\|$. Hence $0 \leqq f(x) \leqq K$ for all $x \in G$. Hence the above considerations give (1) for $\frac{1}{K}f$—and consequently for f too.

Finally drop all extra assumptions. Put

$$f'(x) = \frac{1}{2}(|f(x)| + f(x)), \qquad f''(x) = \frac{1}{2}(|f(x)| - f(x)).$$

Then f', f'' fulfill our original assumptions along with f. Since $f'(x), f''(x) \geqq 0$ for all $x \in G$, so (1) holds for f', f''. Since $f = f' - f''$, (1) extends to f too.

Thus the proof is completed.

11. Means and measures

Consider the C of §4. We shall consider measures $\mu(M)$ in G with the usual properties of *general Lebesgue measure*, including *regularity*. For such a measure $\sigma(M)$ the integral

$$(1) \qquad \int_G f(x)\, d\sigma(x)$$

can be formed for the well-known class of *summable* functions. (Cf. S. Saks, *Theory of the integral*, Warszawa, 1937, p. 19 et seq.)

In particular every $f \in \mathcal{F}_C$ is summable: $f(x)$ is continuous and bounded for $x \in C$, it vanishes for $x \notin C$ and $\mu(C)$ is finite. (Cf. §8.) Accordingly form the expression

$$(2) \qquad \mathbb{M}(f) = \int_G f(x)\, d\sigma(x) \quad \text{for all } f \in \mathcal{F}_C.$$

This $\mathbb{M}(f)$ is obviously a *mean* in the sense of the definition of §9.

By (1) in §8 $f \in \mathcal{F}_C$ implies $f \in \mathcal{F}_{C^i}$ so $f(x) = 0$ for $x \notin C^i$. Consequently it seems natural to restrict the study of (2) to measures with

$$(3) \qquad \sigma(\widetilde{C^i}) = 0.$$

We define accordingly:

DEFINITION I. A measure $\sigma(M)$ is a C^i*-measure* if it is regular and fulfills the condition (3) above.

LEMMA I. *For every* (*regular*) *measure* $\tau(M)$ *there exists one and only one* C^i*-measure* $\sigma(M)$ *such that*

$$(4) \qquad \sigma(M) = \tau(M) \quad \textit{for all } M \leqq C^i.$$

Thus $\sigma(M)$ *is defined by*

$$(5) \qquad \sigma(M) = \tau(C^i M),$$

and we will call $\sigma(M)$ *the* C^i-piece *of* $\tau(M)$.

PROOF. Obviously (5) defines a (regular) measure $\sigma(M)$ and it implies (3), (4). Conversely: Let $\sigma(M)$ fulfill (3), (4). Then $\sigma(-C^i) = 0$ and $(-C^i)M \leqq -C^i$ give $\sigma((-C^i)M) = 0$. Now

$$\sigma(M) = \sigma(C^i M) + \sigma((-C^i)M) = \sigma(C^i M) = \tau(C^i M),$$

proving (5).

(2) establishes a (one to possibly many) correspondence between (possibly not all) means $\mathbb{M}(f)$ and (all) C^i-measures $\sigma(M)$. We now proceed to show that this correspondence is one to one and that it covers all means.

LEMMA II. *Given a mean* $\mathbb{M}(f)$ *there exists at most one* C^i*-measure* $\sigma(M)$ *to which it corresponds by* (2) *above.*

PROOF. Let $\mathbb{M}(f)$ correspond to two C^i-measures $\sigma_1(M)$ and $\sigma_2(M)$ by (2). We must prove that $\sigma_1(M)$ and $\sigma_2(M)$ are identical. I.e. $\sigma_1(M) = \sigma_2(M)$ for all Borel sets M. Since $\sigma_1(M), \sigma_2(M)$ are both regular it suffices to prove $\sigma_1(O) = \sigma_2(O)$ for all O. Now $\sigma_1(O) = \sigma_1(O) = \sigma_1(O \cdot C^i)$, $\sigma_2(O) = \sigma_2(O \cdot C^i)$ (apply (5) in Lemma I with σ_1, σ_2 and with σ_2, σ_2 in place of its σ, τ). Hence we may replace O by $O \cdot C^i$ in our preceding equation—i.e. we may assume $O \leqq C^i$. I.e.: We must only prove

$$\sigma_1(O) = \sigma_2(O) \quad \text{if } O \leqq C^i. \tag{6}$$

Consider a $D \leqq O$. Then there exists an everywhere continuous function $f(x)$ with

(i) $f(x) = 1$ for $x \in D$,

(ii) $f(x) = 0$ for $x \notin O$,

(iii) $O \leqq f(x) \leqq 1$ for all $x \in G$.[2] Then application of (2) with $\sigma_1(M)$ and with $\sigma_2(M)$ gives

$$\sigma_1(D) \leqq \int_G f(x)\, d\sigma_1(x) = \mathbb{M}(f),$$

$$\sigma_2(D) \geqq \int_G f(x)\, d\sigma_2(x) = \mathbb{M}(f);$$

hence

$$\sigma_1(D) \leqq \sigma_2(O). \tag{7}$$

Since (7) holds for all $D \leqq O$, the regularity of σ_1 gives

$$\sigma_1(O) \leqq \sigma_2(O). \tag{8}$$

Interchanging of $\sigma_1(M), \sigma_2(M)$ transforms (8) into (6) as desired.

LEMMA III. *Given a mean* $\mathbb{M}(f)$ *there exists at least one* C^i*-measure* $\sigma(M)$ *which corresponds to it by* (2) *above.*

PROOF. Let $\mathbb{M}(f)$ be given; we will construct such a $\sigma(M)$ explicitly.

Consider an arbitrary O and define

$$\rho(O) = \sup_{f \in \overline{\mathcal{F}}_O, f \in \mathcal{F}_C} \mathbb{M}(f). \tag{9}$$

(The requirement $f \in \overline{\mathcal{F}}_O, f \in \mathcal{F}_C$ is clearly equivalent to $f \in \overline{\mathcal{F}}_{O \cdot C}$. And, since we may replace C by C^i—cf. (1) in §8—we may also write $f \in \overline{\mathcal{F}}_{O \cdot C^i}$. So we have

$$\rho(O) = \sup_{f \in \overline{\mathcal{F}}_{O \cdot C}} \mathbb{M}(f) = \sup_{f \in \overline{\mathcal{F}}_{O \cdot C^i}} \mathbb{M}(f).) \tag{9'}$$

[2]This is Urysohn's lemma. It was first stated in P. Urysohn, *Über die Mächtigkeit der zusammenhängenden Mengen*, Math. Ann., vol. 94 (1925), p. 262 (particularly pp. 290–291). See also P. Urysohn, *Zum Metrisationsproblem*, Math. Ann., vol. 94 (1925), p. 309 (particularly pp. 310–311).

$\rho(O)$ is obviously finite (cf. Lemma II in §9) and $\geqq 0$. Define further for every D

(10) $$\lambda(D) = \inf_{O \geqq D} \rho(O).$$

We derive now various properties of these set functions $\rho(O)$ and $\lambda(D)$.

[A] $$\rho(O) = \sup_{D \leqq O} \lambda(D).$$

Indeed: (10) applies $\rho(O) \geqq \lambda(D)$ whenever $O \geqq D$; hence $\rho(O) \geqq \sup_{D \leqq O} \lambda(D)$. Thus we need only prove $\rho(O) \leqq \sup_{D \leqq O} \lambda(D)$; i.e., considering (9)

(11) $$\mathbb{M}(f) \leqq \sup_{D \leqq O} \lambda(D) \text{ if } f \in \overline{\mathcal{F}}_O, f \in \mathcal{F}_C.$$

Consider accordingly such an f. For any $\varepsilon > 0$ form the set D_ε of all x with $f(x) \geqq \varepsilon$. Since f is continuous, D_ε is closed, since $f \in \mathcal{F}_C$, so $D_\varepsilon \leqq C$. Hence D_ε is compact along with C. Besides $f \in \overline{\mathcal{F}}_O$ implies $D_\varepsilon \leqq O$.

Define f_ε by

$$f_\varepsilon(x) = \text{Max}(f(x) - \varepsilon, O).$$

Always $f_\varepsilon(x) \geqq 0$, always $f_\varepsilon(x) \leqq f(x) \leqq 1$, and $x \notin D_\varepsilon$ implies $f(x) < \varepsilon$, $f_\varepsilon(x) = 0$. So $f_\varepsilon \in \overline{\mathcal{F}}_{D_\varepsilon}$. Thus $f_\varepsilon \in \mathcal{F}_C$, and for every $P \geqq D_\varepsilon$, $f_\varepsilon \in \overline{\mathcal{F}}_P$. Consequently application of (9) to P gives

$$\rho(P) \geqq \mathbb{M}(f_\varepsilon) \quad \text{if } P \geqq D_\varepsilon.$$

Hence application of (10) to D_ε gives

$$\lambda(D_\varepsilon) \geqq \mathbb{M}(f_\varepsilon),$$

and this implies, since $D_\varepsilon \leqq O$,

(12) $$\mathbb{M}(f_\varepsilon) \leqq \sup_{D \leqq O} \lambda(D).$$

Further

$$f(x) - f_\varepsilon(x) = \text{Min}(\varepsilon, f(x)) \left\{ \begin{array}{l} \leqq \varepsilon \\ \geqq 0 \end{array} \right| ,$$

so $\|f - f_\varepsilon\| \leqq \varepsilon$. Hence

$$|\mathbb{M}(f) - \mathbb{M}(f_\varepsilon)| = |\mathbb{M}(f - f_\varepsilon)| \leqq |||\mathbb{M}||| \cdot \|f - f_\varepsilon\| \leqq |||\mathbb{M}|||\varepsilon,$$

(13) $$\mathbb{M}(f_\varepsilon) \geqq \mathbb{M}(f) - |||\mathbb{M}|||\varepsilon.$$

Combination of (12), (13) gives

(14) $$\mathbb{M}(f) - |||\mathbb{M}|||\varepsilon \leqq \sup_{D \leqq O} \lambda(D).$$

Since (14) holds for all $\varepsilon > 0$ it yields (11) and thereby proves [A].

[B] $$\rho(O) + \rho(P) \leqq \rho(Q) \quad \text{if } O \cdot P = \theta \text{ and } O + D \leqq Q.$$

Indeed: Assume $f \in \overline{\mathcal{F}}_O$ and $\overline{\mathcal{F}}_C$, $g \in \overline{\mathcal{F}}_P$ and $\mathcal{F}_C$; then $f + g \in \overline{\mathcal{F}}_Q$ (owing to $O \cdot P = \theta$ and $O + P \leqq Q$). Hence (use (9))

$$\mathbb{M}(f) + M(g) = M(f + g) \leqq \rho(Q),$$

and passing to the $\sup_f$ and $\sup_g$ (use again (9)) we obtain [B].

[C] $$\rho(O) + \rho(P) \geqq \lambda(D) \quad \text{if } O + P \geqq D.$$

Indeed: G is locally compact; consequently there exists a $Q \geqq D$ such that $\overline{Q}$ is compact and $\leqq O + P$. Now by (10) $\rho(Q) \geqq \lambda(D)$; hence we need only prove

$$\rho(O) + \rho(P) \geqq \rho(Q). \tag{15}$$

Owing to (9) this amounts to showing

$$\rho(O) + \rho(P) \geqq \mathbb{M}(f) \quad \text{if } f \in \overline{\mathcal{F}}_Q, f \in \mathcal{F}_C. \tag{16}$$

Consider accordingly such an f. $\widetilde{O}\overline{Q}, \widetilde{P}Q$ are closed sets, they are $\leqq \overline{Q}$ which is compact, and hence they are compact. They are disjoint since $\widetilde{O}\overline{Q} \cdot \widetilde{P} \cdot \overline{Q} = \widetilde{O}\widetilde{P}\overline{Q} = (\widetilde{O + P})\overline{Q} = \theta$ owing to $\overline{Q} \leqq O + P$. So $\widetilde{P}\overline{Q}$ is closed, $\widetilde{\widetilde{O}\overline{Q}}$ is open, and $\widetilde{P}\overline{Q} \leqq \widetilde{\widetilde{O}\overline{Q}}$. Then there exists an everywhere continuous function $h(x)$ with

(i) $h(x) = 1$ for $x \in \widetilde{P}\overline{Q}$,

(ii) $h(x) = 0$ for $x \in \widetilde{O}\overline{Q}$ (i.e. $x \notin \widetilde{\widetilde{O}\overline{Q}}$),

(iii) $0 \leqq h(x) \leqq 1$ for all $x \in G$.

(Cf. footnote 2 on p. 104. Our $\widetilde{P}, \overline{Q}, \widetilde{\widetilde{Q}\overline{Q}}$ stand for the D, O there.) Now form

$$k(x) = f(x)h(x), \qquad l(x) = f(x)(1 - h(x)).$$

Then clearly $k, l \in \mathcal{F}_C$ along with $f \in \mathcal{F}_C$, and always $0 \leqq k(x), l(x) \leqq 1 \cdot k(x) \neq 0$ implies $f(x) \neq 0$; hence, owing to $f \in \overline{\mathcal{F}}_Q$, $x \in Q$. It implies further $h(x) \neq 0$; hence by (ii) $x \notin \widetilde{O}\overline{Q}$. But $x \in Q \leqq \overline{Q}$; consequently $x \notin \widetilde{O}$, $x \in O$. So $k \in \overline{\mathcal{F}}_O$. Similarly, by (i) $l \in \overline{\mathcal{F}}_P$. Now

$$\mathbb{M}(f) = \mathbb{M}(k + l) = \mathbb{M}(k) + \mathbb{M}(l) \leqq \rho(O) + \rho(P),$$

proving (16) and consequently [C].

$$\lambda(D) + \lambda(E) \geqq (D + E). \tag{D}$$

Indeed: For $O \geqq D$, $P \geqq E$ we have $O + P \geqq D + E$, so [C] (with our $D + E$ in place of its D) gives

$$\rho(O) + \varphi(P) \geqq \lambda(D + E).$$

Passing to the $\sup_O$ and $\sup_P$ (use (10)) [D] obtains.

$$\lambda(D) + \lambda(E) = \lambda(D + E) \quad \text{if } D \cdot E = \theta. \tag{E}$$

Indeed: By [D] it suffices to prove

$$\lambda(D) + \lambda(E) \leqq \lambda(D + E). \tag{17}$$

Owing to (10) this amounts to showing

$$\lambda(D) + \lambda(E) \leqq \rho(O) \quad \text{if } O \geqq D + E. \tag{18}$$

G is locally compact; consequently $D \cdot E = \theta$ implies the existence of two $P \geqq D$, $Q \geqq E$ with $P \cdot Q = \theta$. Replacement of P, Q by $P \cdot O$, QO does not affect these properties; i.e. we may assume $P, Q \leqq O$. Now $D \leqq P$, $E \leqq Q$ give $\lambda(D) \leqq \rho(P), \lambda(E) \leqq \rho(Q)$ by (10). Next $P \cdot Q = \theta$, $P + Q \leqq O$ give $\rho(P) + \rho(Q) \leqq \rho(O)$. Combining these inequalities gives (18) and thereby proves [E].

Since $\lambda(D)$ is always $\geqq 0$ and finite and since it possesses the properties [D] and [E] above, it follows that it fulfills all requirements of §2, Chapter I. We can therefore use it to define with its help the $\mu(O)$ and $\nu(M)$ (any Borel set $\leqq G$), following the procedure of §2, Chapter I. Comparison of [A] above with (2.1.4), Chapter I, shows that the $\mu(O)$ thus obtained coincides with our $\rho(O)$. We also know from (2.3.2), Chapter I, that $\mu(O) = \nu(O)$; hence we have

(19) $$\rho(O) = \mu(O) = \nu(O).$$

The regularity of $\rho(O)$ gives $\nu(D) = \inf_{O \geqq D} \rho(O)$. (Cf. (4.2), Chapter I, and (18.5.2), Chapter IV.) This, together with (10), yields

(20) $$\lambda(D) = \nu(D).$$

Since the $\mu(O)$ coincides with our $\rho(O)$, we prefer to denote the resulting $\nu(M)$ by $\sigma(M)$.

We claim now:

[F] $\sigma(M)$ is a C^i-measure (cf. Definition I, above), and $\sigma(D) = \rho(O)$ (cf. (9)), $\sigma(D) = \lambda(D)$ (cf. (10)).

Indeed: $\sigma(M)$ is a regular Lebesgue measure (i.e. it fulfills the conditions of §18, Chapter IV) by its nature. The two equations are just restatements of (19) and (20). Thus we must only prove (3) (cf. Definition I above).

Now C^i, G are both open sets and it is clear from (9′) that $\rho(C^i) = \rho(G)$ and is finite. By the equations already established (i.e. by (19)) this means that $\sigma(C^i) = \sigma(G)$ and is finite. Consequently $\sigma(\widetilde{C^i}) = \sigma(G) - \sigma(C^i) = 0$ which coincides with (3) and thereby proves [F].

Having thus obtained the desired $\sigma(M)$ we proceed to establish the correspondence (2) in several steps.

[G] $$\mathbb{M}(f) \leqq \int_G f(x)\, d\sigma(x) \quad \text{for all } f \in \overline{\mathcal{F}}_C.$$

Indeed: Consider an $f \in \overline{\mathcal{F}}_C$ and an $\varepsilon > 0$. Choose $n = 1, 2, \ldots$ with $n\varepsilon \geqq 1$. Define $f_i(x) = \text{Min}\ (i\varepsilon, f(x))$ for $i = 0, 1, \ldots, n-1, n$. Then clearly all $f_i \in \overline{\mathcal{F}}_C$ along with $f \in \overline{\mathcal{F}}_C$ and also

(21) $$0 = f_0(x) \leqq f_1(x) \cdots \leqq f_{n-1}(x) \leqq f_n(x) = f(x),$$

and

(22) $$f_i(x) - f_{i-1}(x) = \begin{cases} \varepsilon & \text{for } f(x) > i\varepsilon, \\ & f(x) - (i-1)\varepsilon \text{ for } (i-1)\varepsilon < f(x) \leqq i\varepsilon, \\ 0 & \text{for } f(x) \leqq (i-1)\varepsilon \\ & \text{and } i = 1, \ldots, n-1, n. \end{cases}$$

Define further

(23) $$g_i(x) = \frac{1}{\varepsilon}(f_i(x) - f_{i-1}(x)) \quad \text{for } i = 1, \ldots, n-1, n$$

and

(24) $$O_i = \text{ set of all } x \text{ with } f(x) > i\varepsilon \text{ for } i = 0, 1, \ldots, n-1, n.$$

Now (21), (22) give

(25) $$f(x) = \varepsilon \sum_{i=1}^{n} g_i(x).$$

(22), (23) show that always $0 \leqq g_i(x) \leqq 1$, and (22), (23), (24) show that $g_i(x) = 0$ for $x \notin O_{i-1}$. Hence $g_i \in \overline{\mathcal{F}}_{O_{i-1}}$. Clearly all $g_i \in \mathcal{F}_C$ because all $f_i \in \mathcal{F}_C$. Consequently (9) and [F] give

(26) $$\mathbb{M}(g_i) \leqq \rho(O_{i-1}) = \sigma(O_{i-1}).$$

Finally (24) makes it clear that

(27) $$O_0 \geqq O_1 \geqq \cdots \geqq O_{n-1} \geqq O_n = \theta.$$

Now by (25), (26)

(28) $$\mathbb{M}(f) = \varepsilon \sum_{i=1}^{n} \mathbb{M}(g_i) \leqq \varepsilon \sum_{i=1}^{n} \sigma(O_{i-1}).$$

But (27) gives

$$\sigma(O_{i-1}) = \sigma(O_{i-1}) - \sigma(O_n) = \sum_{j=i}^{n} \{\sigma(O_{j-1}) - \sigma(O_j)\}$$
$$= \sum_{j=1}^{n} \sigma(O_{j-i} \cdot \widetilde{O}_j),$$
$$\sum_{i=1}^{n} \sigma(O_i) = \sum_{j=1}^{n} (j+1)\sigma(O_{j-i} \cdot \widetilde{O}_j).$$

Hence (28) becomes

(29) $$M(f) \leqq \sum_{j=1}^{n} (j+1)\varepsilon\sigma(O_{j-1} \cdot \widetilde{O}_j).$$

On the other hand since the $O_{j-1} \cdot \widetilde{O}_j$, $j = 1, \ldots, n$, are disjoint by (27), therefore

$$\int_Q (f(x) + \varepsilon)\, d\sigma(x) \geqq \sum_{j=1}^{n} \int_{O_{j-1}\widetilde{O}_j} (f(x) + \varepsilon)\, d\sigma(x) \geqq \sum_{j=1}^{n} \int_{O_{j-1}\widetilde{O}_j} (j+1)\varepsilon\, d\sigma(x)$$
$$= \sum_{j=1}^{n} (j+1)\varepsilon\sigma(O_{j-1} \cdot \widetilde{O}_j),$$

i.e.

(30) $$\int_G (f(x) + \varepsilon)\, d\sigma(x) \geqq \sum_{j=1}^{n} (j+1)\varepsilon\sigma(O_{j-1}\widetilde{O}_j).$$

(29), (20) combine to

(31) $$\mathbb{M}(f) = \int_G (f(x) + \varepsilon)\, d\sigma(x);$$

since $\sigma(G) = \rho(G)$ is finite (cf. in [F] or before (10)), the validity of (31) for all $\varepsilon > 0$ implies [G].

[H] $$\mathbb{M}(f) = \int_G f(x)\, d\sigma(x) \quad \text{for all } f \in \overline{\mathcal{F}}_C.$$

Indeed: Consider an $f \in \overline{\mathcal{F}}_C$ and an $\varepsilon > 0$.

Application of (9′) with $O = G$ permits us to choose an $h \in \overline{\mathcal{F}}_C$ with

$$M(h) \geqq \sigma(G) - \varepsilon. \tag{32}$$

Define $g(x) = \operatorname{Max}(O, h(x) - f(x))$; then $g \in \overline{\mathcal{F}}_C$ along with $f, h \in \mathcal{F}_C$ and clearly

$$f(x) + g(x) = \operatorname{Max}(f(x), h(x)) \left\{ \begin{array}{l} \geqq h(x) \\ \leqq 1 \end{array} \right| . \tag{33}$$

(32), (33) give

$$\mathbb{M}(f) + \mathbb{M}(g) = \mathbb{M}(f+g) \geqq \mathbb{M}(h) \geqq \sigma(G) - \varepsilon,$$

$$\begin{aligned} \int_G f(x)\,d\sigma(x) + \int_G g(x)\,d\sigma(x) &= \int_G (f(x) + g(x))\,d\sigma(x) \\ &\leqq \int_G d\sigma(x) = \sigma(G); \end{aligned}$$

hence

$$\mathbb{M}(f) + \mathbb{M}(g) \geqq \int_G f(x)\,d\sigma(x) + \int_G g(x)\,d\sigma(x) - \varepsilon. \tag{34}$$

On the other hand application of [G] to f and g gives

$$\mathbb{M}(f) \leqq \int_G f(x)\,d\sigma(x), \qquad \mathbb{M}(g) \leqq \int_G g(x)\,d\sigma(x). \tag{35}$$

Combining (34), (35) we obtain

$$\int_G f(x)\,d\sigma(x) - \varepsilon \leqq \mathbb{M}(f) \leqq \int_G f(x)\,d\sigma(x). \tag{36}$$

Since this is true for all $\varepsilon > 0$ it implies [H].

[I] $$\mathbb{M}(f) = \int_G f(x)\,d\sigma(x) \quad \text{for all } f \in \mathcal{F}_C.$$

Indeed: Consider an $f \in \mathcal{F}_C$. Assume first $f(x) \geqq 0$ for all $x \in G$. Choose a $A > 0$ with $K \geqq \|f\|$. Hence $0 \leqq f(x) \leqq K$ for all $x \in G$. Then $\frac{1}{K} f \in \overline{\mathcal{F}}_C$, so [H] applies to $\frac{1}{K} f$. I.e. our [I] holds for $\frac{1}{K} f$—and consequently for f too.

Now drop the extra assumption. Put

$$f'(x) = \frac{1}{2}(|f(x)| + f(x)), \qquad f''(x) = \frac{1}{2}(|f(x)| - f(x)).$$

Then $f', f'' \in \mathcal{F}_C$ along with $f \in \mathcal{F}_C$. Since $f'(x), f''(x) \geqq 0$ for all $x \in G$, so [I] holds for f', f''. Since $f = f' - f''$ [I] extends to f too.

[F] and [I] together complete the proof.

The definition which follows sums up the situation.

Definition II. A mean $\mathbb{M}(f)$ and a C^i-measure $\sigma(M)$ which correspond to each other by (2) will briefly be called *correspondents*. By Lemma II and III this is a one to one correspondence of all means with all C^i-measures.

We conclude with this observation:

Lemma IV. *If* $\mathbb{M}(f)$ *and* $\sigma(M)$ *are correspondents, then*

$$|||M||| = \sigma(C^i) = \sigma(C) = \sigma(G).$$

PROOF. By (9′) in the proof of Lemma III above

$$\rho(C^i) = \rho(G) = \sup_{f \in \overline{\mathcal{F}}_{C^i}} \mathbb{M}(f). \tag{37}$$

By (19) in the same proof the two left-hand side terms are equal to $\sigma(C^i)$ and to $\sigma(G)$ respectively. By (1) in §8 the right-hand side term is equal to $\sup_{f \in \overline{\mathcal{F}}_C} \mathbb{M}(f)$, and this is by (η'') in Lemma II in §9 equal to $|||\mathbb{M}|||$ so (37) becomes

$$\sigma(C^i) = \sigma(G) = |||\mathbb{M}|||. \tag{38}$$

(38) gives the first one and the third one of the asserted equations. The second one ensues, owing to $\sigma(C^i) \leqq \sigma(C) \leqq \sigma(G)$.

12. Left invariance of measures

Consider the C of §4. We assume now $C^i \neq \theta$.

The correspondence of Definition II in §11 will be specialized in this section to the case when the mean $\mathbb{M}(f)$ is *left invariant* (abbreviated: l.i.) in the sense of Definition I in §10. We wish to investigate the corresponding C^i-measures $\sigma(M)$.

The pertinent definitions for measures are these:

DEFINITION I. A measure $\tau(M)$ (i.e. a regular measure in G) is *left invariant* (abbreviated: l.i.) if

$$\tau(aM) = \tau(M) \quad \text{for all } a \in G \text{ and } M \leqq G.$$

DEFINITION II. A C^i-measure $\sigma(M)$ is C^i-*left invariant* (abbreviated: C^i-l.i.) if

$$\sigma(aM) = \sigma(M) \quad \text{whenever } M, aM \leqq C^i.$$

The lemmas which follow determine the connection between these two notions.

LEMMA I. *If $\tau(M)$ is an l.i. measure, then its C^i-piece $\sigma(M)$ (cf. Lemma* I *in* §11) *is a C^i-l.i. C^i-measure.*

PROOF. If $M, AM \leqq C^i$, then

$$\sigma(M) = \tau(M) = \tau(aM) = \sigma(aM).$$

LEMMA II. *Given a C^i-l.i. C^i-measure $\sigma(M)$, there exists at most one l.i. measure $\tau(M)$ the C^i-piece of which (cf. above) is $\sigma(M)$.*

PROOF. Let $\sigma(M)$ be the C^i-piece of both l.i. measures $\tau_1(M)$ and $\tau_2(M)$. Then $M \leqq C^i$ implies $\tau_1(M) = \sigma(M) = \tau_2(M)$, and $M \leqq aC$ implies $a^{-1}M \leqq C^i$,

$$\tau_1(M) = \tau_1(a^{-1}M) = \tau_2(a^{-1}M) = \tau_2(M).$$

So we have

$$\tau_1(M) = \tau_2(M) \quad \text{whenever } M \leqq aC^i. \tag{1}$$

Consider now an arbitrary D. Then $D \leqq \sum_{i=1}^{k} a_i \cdot C^i$. (This is due to $D \leqq \sum_a aC^i$, cf. the arguments made in §8. The above relation makes use of $C^i \neq \theta$.) Consequently,

$$D = \sum_{i=1}^{k} N_i \quad \text{with } N_iN_j = \theta \text{ for } i \neq j \text{ and } N_i \leqq a_i \cdot C^i. \tag{2}$$

(Put $M_i = D \cdot a_i \cdot C^i$, $N_i = M_i(\sum_{j=1}^{(i-1)} M_j)$.) Now (1), (2) give

$$\tau_1(D) = \sum_{i=1}^{k} \tau_i(N_i) = \sum_{i=1}^{k} \tau_2(N_i) = \tau_2(D),$$

i.e.

$$\tau_1(D) = \tau_2(D). \tag{3}$$

Since $\tau_1(M), \tau_2(M)$ are regular, (3) establishes $\tau_1(M) = \tau_2(M)$ for all Borel sets M, i.e. the identity of $\tau_1(M)$ and $\tau_2(M)$.

LEMMA III. *Given C^i-l.i. C^i-measure $\sigma(M)$, there exists at least one l.i. measure $\tau(M)$ the C^i-piece of which (cf. above) is $\sigma(M)$.*

PROOF. Let $\sigma(M)$ be given; we will construct such $\tau(M)$ explicitly.

Consider an arbitrary D. As in (2) in the proof of Lemma II, we can put this D into the form

$$D = \sum_{i=1}^{k} N_i \quad \text{with } N_i N_j = \theta \text{ for } i \neq j, \text{ and } N_i \leqq a_i C^i \tag{4}$$

$$(k = 1, 2, \ldots, a_1, \ldots, a_k \in G).$$

(Again $C^i \neq \theta$ is used, cf. loc. cit.) Thus $a_i^{-1} N_i \leqq C^i$, and so we can form

$$\lambda(D) = \sum_{i=1}^{k} \sigma(a_i^{-1} N_i). \tag{5}$$

We derive now various properties of this expression $\lambda(D)$.

[A] $\lambda(D)$ is a function of D only. I.e. it is independent of the particular decomposition (4) of D used.

Indeed: Consider two such decompositions

$$D = \sum_{i=1}^{k} N_i \quad \text{with } N_i N_j = \theta \text{ for } i \neq j \text{ and } N_i \leqq a_i C^i,$$

$$D = \sum_{p=1}^{k'} N'_p \quad \text{with } N'_p N'_q = \theta \text{ for } p \neq q \text{ and } N'_p \leqq a'_p C^i.$$

Then $N_i \leqq D = \sum_{p=1}^{k'} N'_p$; hence $N_i = \sum_{p=1}^{k'} N_i N'_p$, $a_i^{-1} N_i = \sum_{p=1}^{k'} a_i^{-1}(N_i N'_p)$ and so

$$\sum_{i=1}^{k} \sigma(a_i^{-1} N_i) = \sum_{i=1}^{k} \sum_{p=1}^{k'} \sigma(a_i^{-1}(N_i N'_p)). \tag{6}$$

Similarly

$$\sum_{p=1}^{k'} \sigma(a_i'^{-1} N'_p) = \sum_{i=1}^{k} \sum_{p=1}^{k'} \sigma(a_i^{-1}(N_i N'_p)). \tag{7}$$

Now $a_i^{-1}(N_i N'_p) \leqq C^i$, $a_p'^{-1}(N_i N'_p) \leqq C^i$ and $a_p'^{-1}(N_i N'_p) = (a_p'^{-1} a_i)(a_i^{-1}(N_i N'_p))$. Consequently the C^i-l.i. of $\sigma(M)$ gives

$$\sigma(a_i^{-1}(N_i N'_p)) = \sigma(a_p'^{-1}(N_i N'_p)). \tag{8}$$

Combining (6), (7), (8) we obtain

$$\sum_{i=1}^{k} \sigma(a_i^{-1} N_i) = \sum_{p=1}^{k'} \sigma(a_p'^{-1} N_p'), \tag{9}$$

proving [A].

[B] $$\lambda(D) + \lambda(E) \geqq \lambda(D+E).$$

[C] $$\lambda(D) + \lambda(E) = \lambda(D+E) \quad \text{if } D \cdot E = \theta.$$

Indeed: Apply (4) to $D + E$:

$$D + E = \sum_{i=1}^{k} N_i \quad \text{with } N_i N_j = \theta \text{ for } j \neq i \text{ and } N_i \leqq a_i C^i. \tag{10}$$

Then

$$\left.\begin{aligned} D &= \textstyle\sum_{i=1}^{k} N_i D \qquad (11) \\ E &= \textstyle\sum_{i=1}^{k} N_i E \qquad (12) \end{aligned}\right\} \quad \text{with the same } k \text{ and } a_i, \ldots, a_k,$$

can be used as (4) for D and for E respectively. Consequently we have

$$\begin{aligned} \lambda(D+E) &= \sum_{i=1}^{k} \sigma(a_i^{-1} N_i), \\ \lambda(D) &= \sum_{i=1}^{k} \sigma(a_i^{-1}(N_i D)), \\ \lambda(E) &= \sum_{i=1}^{k} \sigma(a_i^{-1}(N_i E)). \end{aligned} \tag{13}$$

Besides $N_i \leqq D + E$, so $N_i = N_i D + N_i E$, $a_i^{-1} N_i = a_i^{-1}(N_i D) + a_i^{-1}(N_i E)$; hence

$$\sigma(a_i^{-1}(N_i D)) + \sigma(a_i^{-1}(N_i E)) \geqq \sigma(a_i^{-1} N_i), \tag{14}$$

$$\sigma(a_i^{-1}(N_i D)) + \sigma(a_i^{-1}(N_i E)) = \sigma(a_i^{-1} N_i) \quad \text{if } D \cdot E = \theta. \tag{15}$$

Clearly (13), (14) give [B] and (13), (15) give [C].

Since $\lambda(D)$ is always $\geqq 0$ and finite and since it possesses the properties [B] and [C] above, it follows that it fulfills the requirements of §2, Chapter I. We can therefore use it to define with its help the $\mu(O)$ and $\nu(M)$ (M any Borel set $\leqq G$), following the procedure of §2, Chapter I. We prefer to denote the resulting $\nu(M)$ by $\tau(M)$.

We continue our discussion.

[D] $\tau(M)$ is an l.i. measure.

Indeed: We must prove

$$\tau(aM) = \tau(M).$$

Considering how $\tau(M)$ is defined with the help of $\lambda(D)$, it suffices to prove

$$\lambda(aD) = \lambda(D). \tag{16}$$

Inspection of (4) shows that it remains unaffected when we replace in it D, a_i, N_i by aD, aa_i, aN_i. Since $(aa_i)^{-1}(aN_i) = a_i^{-1}N_i$, (5) now shows that $\lambda(aD) = \lambda(D)$, i.e. that (16) is true.

[E] $$\lambda(D) = \sigma(D) \quad \text{if } D \leqq C^i.$$

Indeed: If $D \leqq C^i$, then we can choose in (4) $k = 1$ and $a_1 = 1, N_1 = D$. Then (5) yields [E].

[F] $$\sigma(M) \text{ is the } C^i\text{-piece of } \tau(M).$$

Indeed: We must prove

(17) $$\tau(M) = \sigma(M) \quad \text{if } M \leqq C^i.$$

It suffices, however, to prove

(18) $$\tau(O) = \sigma(O) \quad \text{if } O \leqq C^i.$$

Because we may then apply $\inf_{M \leqq O \leqq C^i}$ to both sides of (18), obtaining

(19) $$\inf_{M \leqq O \leqq C^i} \tau(O) = \inf_{M \leqq O \leqq C^i} \sigma(O) \quad \text{if } M \leqq C^i.$$

Now clearly each $\inf_{M \leqq O \leqq C^i}$ in (19) is $\geqq$ the corresponding $\inf_{O \geqq M}$. On the other hand $O \geqq M$ implies $M \leqq O \cdot C^i \leqq C^i$, and $\tau(O \cdot C^i) \leqq \tau(O)$, $\sigma(O \cdot C^i) \leqq \sigma(O)$. So each $\inf_{M \leqq O \leqq C^i}$ in (19) is also $\leqq$ the corresponding $\inf_{O \geqq M}$. Hence we have $=$, and (19) becomes

(20) $$\inf_{O \geqq M} \tau(O) = \inf_{O \geqq M} \sigma(O) \quad \text{if } M \leqq C^i,$$

and since $\tau(M), \sigma(M)$ are regular this implies (18) by (18.2), Chapter IV.

We now prove (18). Consider an $O \leqq C^i$. Apply $\sup_{D \leqq O}$ to [E]; then we obtain

(21) $$\sup_{D \leqq O} \lambda(D) = \sup_{D \leqq O} \sigma(D) \quad \text{if } O \leqq C^i.$$

Owing to the definition of $\mu(O)$, the left-hand side of (21) is $\mu(O)$, and this is equal to $\tau(O)$. Since $\sigma(M)$ is regular, the right-hand side of (21) is $\sigma(O)$. Hence (21) gives $\tau(O) = \sigma(O)$, i.e. (18), as desired.

[D] and [F] together complete the proof.

The definition which follows sums up the situation.

DEFINITION III. An l.i. measure $\tau(M)$ and a C^i-l.i. C^i-measure $\sigma(M)$, where $\sigma(M)$ is the C^i-piece of $\tau(M)$, will briefly be called *correspondents.*

By Lemmas I and II this is a one to one correspondence of all l.i. measures with all C^i-l.i. C^i-measures.

13. Means and measures (concluded)

Consider the C of §4, assuming $C^i \neq \theta$.

We establish the connection between our various notions of left invariance for means on one hand and for measures on the other hand.

LEMMA I. *Let the mean* $\mathbb{M}(f)$ *and the* C^i*-measure* $\sigma(M)$ *be correspondents* (*in the sense of Definition* II *in* §11). *Then* $\mathbb{M}(f)$ *is l.i. if and only if* $\sigma(M)$ *is* C^i*-l.i.*

PROOF. That $\mathbb{M}(f)$ is l.i. means that we have for every $a \in G$ this property:

(1) $\mathbb{M}_x(f(ax)) = \mathbb{M}_x(f(x))$ whenever $f(ax), f(x)$ belong both to $\mathcal{F}_C$.

That $f(ax)$ belongs to $\mathcal{F}_C$ means that $f(x)$ belongs to $\mathcal{F}_{a^{-1}C}$. Hence the hypothesis of (1) means that $f(x)$ belongs to $\mathcal{F}_{C^i}$ and to $\mathcal{F}_{a^{-1}C}$, i.e. that it belongs to $\mathcal{F}_{C\cdot a^{-1}C}$. Now define

(2) $$\mathbb{M}'_x(f(x)) = \mathbb{M}_x(f(x)) \quad \text{for the } f(x) \text{ of } \mathcal{F}_{C\cdot a^{-1}C},$$

(3) $$\mathbb{M}''_x(f(x)) = \mathbb{M}_x(f(ax)) \quad \text{for the } f(x) \text{ of } \mathcal{F}_{C\cdot a^{-1}C};$$

then $\mathbb{M}', \mathbb{M}''$ are both means for $C \cdot a^{-1}C$ in place of C and (1) amounts to stating the identity of $\mathbb{M}'$ and of $\mathbb{M}''$.

Let $\sigma'(M), \sigma''(M)$ be the $(C \cdot a^{-1}C)^i$-piece—i.e. the $C^i \cdot a^{-1} \cdot C^i$-piece—of $\sigma(M), \sigma(a^{-1}M)$ respectively.

The form of the definitory relation (2) in §11 (remember Definition II in §11) shows that since $\sigma(M)$ is the correspondent of $\mathbb{M}$ therefore $\sigma'(M), \sigma''(M)$ are the correspondents of $\mathbb{M}', \mathbb{M}''$ respectively. Consequently (1), i.e. the identity of $\mathbb{M}', \mathbb{M}''$, is equivalent to the identity of $\sigma'(M), \sigma''(M)$—i.e. to

$$\sigma'(M) = \sigma''(M) \quad \text{for all } M \leqq (C \cdot a^{-1}C)^i.$$

Owing to (4), (5) in Lemma I in §11 this means

$$\sigma(M) = \sigma(aM) \quad \text{for } M \leqq (C \cdot a^{-1}C)^i = C^i \cdot a^{-1}C^i,$$

i.e.

(4) $$\sigma(M) = \sigma(aM) \quad \text{for } M,\ aM \leqq C^i.$$

Thus the l.i. of $\mathbb{M}$ which is equivalent to the validity of (1) for all $a \in G$ amounts to the validity of (4) for all $a \in G$. But this is precisely the statement of the C^i-l.i. of $\sigma(M)$.

We are now in a position to prove:

LEMMA II. *The relationship*

$$\mathbb{M}(f) = \int_G f(x)\, d\tau(x) \quad \textit{for all } f \in \mathcal{F}_C$$

establishes a one to one correspondence between all l.i. means $\mathbb{M}$, *and all l.i. measures* $\tau(M)$.

PROOF. Denote the C^i-piece of $\tau(M)$ by $\sigma(M)$. Then the above relation is equivalent to

(5) $$\mathbb{M}(f) = \int_G f(x)\, d\sigma(x) \quad \text{for all } f \in \mathcal{F}_C$$

(remember (4) in §11). If $\tau(M)$ is l.i., then $\sigma(M)$ is C^i-l.i., and if $\sigma(M)$ is l.i., then $\tau(M)$ can be chosen l.i., and this correspondence between $\tau(M)$ and $\sigma(M)$ is one to one by Definition III in §12. The correspondence (5) between the l.i. $\mathbb{M}(f)$ and the C^i-l.i. $\sigma(M)$ is also one to one by Definition II in §11 together with Lemma I above. Combining these two facts, we see that our original correspondence between the l.i. $\mathbb{M}(f)$ and the l.i. $\tau(M)$ is also one to one.

To conclude, we define again:

Definition. The l.i. mean $\mathbb{M}(f)$ and the l.i. measure $\mu(M)$ which are correlated by the one to one correspondence of Lemma II above will again be called correspondents.

14. Convergent systems of a.l.i. means

Our ultimate goal is the construction of an l.i. measure. We know from the definition at the end of §13 that this is equivalent to the following proposition:

Choose the C of §4 assuming $C^i \neq \theta$. Then find an l.i. mean $\mathbb{M}(f)$ (in C).

The appropriate tool for this program is contained in the lemma which follows. In connection with this lemma, we wish to observe that it is a simple convergence statement, a detailed proof of which may seem pedantic. We give it nevertheless, because the convergence occurs in an unusual way: With the O, $1 \in O$, as indices. (Cf., however, the ideal $\mathcal{T}$ defined in (15.4), Chapter III. The notion of convergence modulo $\mathcal{T}$ which could be defined in the sense of (11.3), Chapter II, is very nearly what we use here.)

LEMMA I. *Assume that for every O with $1 \in O$ a set $\mathcal{M}_O$, with the properties which follow is given.*

(α) *$\mathcal{M}_O$ is a nonempty set of O-a.l.i. means $\mathbb{M}(f)$ (in C).*

(β) *For every $f \in \overline{\mathcal{F}}_C$ and every $\varepsilon > 0$ there exists a $O_0 = O_0(f, \varepsilon)$ with $1 \in O_O$ such that*

$$1 \in O', \qquad O'' \leqq O_0$$

and

$$\mathbb{M}'(f) \quad \textit{in } \mathcal{M}_{O'}, \qquad \mathbb{M}''(f) \quad \textit{in } \mathcal{M}_{O''}$$

imply $|\mathbb{M}'(f) - \mathbb{M}''(f)| \leqq \varepsilon$.

Then there exists one and only one mean $\overline{\mathbb{M}(f)}$ with the following property:

(γ) *For every $f \in \mathcal{F}_C$ and every $\varepsilon > 0$ there exists an $O_1 = O_1(f, \varepsilon)$ with $1 \in O_1$, such that*

$$1 \in O' \leqq O_1$$

and

$$\overline{\mathbb{M}'}(f) \quad \textit{in } \mathcal{M}_{O'}$$

imply $|\overline{\mathbb{M}(f)} - \mathbb{M}''(f)| \leqq \varepsilon$. (For $f \in \overline{\mathcal{F}}_C$ we can choose the $O_1 = O_1(f, \varepsilon)$ of (γ) as the $O_0 = O_0(f, \varepsilon)$ of (β).)

PROOF. We proceed in several successive steps.

[A] Given an $f \in \overline{\mathcal{F}}_C$ there exists a real number ξ with the following property: If

$$\varepsilon > 0, \quad 1 \in O \leqq O_0(f, \varepsilon), \quad \mathbb{M}(g) \text{ in } \mathcal{M}_O,$$

then

$$|\xi - \mathbb{M}(f)| \leqq \varepsilon.$$

Indeed: For every system $\varepsilon, O, \mathbb{M}(g)$ with

$$\varepsilon > 0, \quad 1 \in O \leqq O_0(f, \varepsilon), \quad \mathbb{M}(g) \text{ in } \mathcal{M}_O, \tag{1}$$

form the (numerical) interval

$$I(\varepsilon, O, \mathbb{M}(g)) : \mathbb{M}(f) - \varepsilon \leqq \xi \leqq \mathbb{M}(f) + \varepsilon. \tag{2}$$

Consider now a finite family of systems $\varepsilon_i, O_i, \mathbb{M}_i(g)$, $i = 1, \ldots, n$, all of which fulfill (1). Consider any two $i, j = 1, \ldots, n$. Form $O^* = O_0(f, \varepsilon_i)O_0(f, \varepsilon_j)$ and choose an $\mathbb{M}^*(g)$ in $\mathcal{M}g^*$. Now apply (β) above to $\varepsilon_i, O_i, O^*, \mathbb{M}_i(g), \mathbb{M}^*(g)$ and to $\varepsilon_j, O_j, O^*, \mathbb{M}_j(g), \mathbb{M}^*(g)$ in place of its $\varepsilon, O', O'', \mathbb{M}'(g), \mathbb{M}''(g)$.

Then

$$|\mathbb{M}_i(f) - \mathbb{M}^*(f)| \leqq \varepsilon_i, \qquad |\mathbb{M}_j(f) - \mathbb{M}^*(f)| \leqq \varepsilon_j$$

result, and therefore

$$\mathbb{M}_i(f) - \varepsilon_i \leqq \mathbb{M}_j(f) + \varepsilon_j. \tag{3}$$

Since (3) holds for all $i, j = 1, \ldots, n$, it implies, in conjunction with (1), that

$$I(\varepsilon_1, O_1, \mathbb{M}_1(g)), \ldots, I(\varepsilon_n, O_n, \mathbb{M}_n(g))) \neq \theta. \tag{4}$$

Since the $I(\varepsilon, O, \mathbb{M}(g))$ are closed, bounded, numerical sets, hence compact sets, it follows from (4) that all $I(\varepsilon, O, \mathbb{M}(g))$ possess a common element ξ_0. (Cf. (5.2), Chapter II.) More precisely: We ought to have closed subsets of a fixed compact set. This can be obtained by picking a fixed system $\varepsilon^+, O^+, \mathbb{M}^+(g)$, which fulfills (1) and then considering the $I(\varepsilon, O, \mathbb{M}(g)) \cdot I(\varepsilon^+, O^+, \mathbb{M}^+(g))$ in place of the $I(\varepsilon, O, \mathbb{M}(g))$. This ξ_0 obviously possesses the property expressed in [A].

[B] Given an $f \in \mathcal{F}_C$, there exists a real number ξ_0 with the following property: There exists for every $\varepsilon > 0$ an $O_1 = O_1(f, \varepsilon)$ such that if

$$\varepsilon > 0, \quad 1 \in O \leqq O_1, \quad \mathbb{M}(g) \text{ in } \mathcal{M}_{O_1},$$

then

$$|\xi_0 - \mathbb{M}(f)| \leqq \varepsilon.$$

Indeed: For $f \in \overline{\mathcal{F}}_C$ [E] follows from [A] with $O_1(f, \varepsilon) = O_0(f, \varepsilon)$. Consider now an $f \in \mathcal{F}_C$. Assume first $f(x) \geqq 0$ for all $x \in G$. Choose a $K > 0$ with $K \geqq \|f\|$. Hence $0 \geqq f(x) \leqq K$, for all $x \in G$. Then $\frac{1}{K}f \in \overline{\mathcal{F}}_C$ so [B] holds for $\frac{1}{K}f$. Consequently it holds for f, too, with $O_1(f, \varepsilon) = O_1(\frac{1}{K}f, \frac{1}{K}\varepsilon)$.

Now drop the extra assumption. Put, as at a previous occasion,

$$f'(x) = \frac{1}{2}(|f(x)| + f(x)), \qquad f''(x) = \frac{1}{2}(|f(x)| - f(x)).$$

Then $f', f'' \in \mathcal{F}_C$ along with $f \in \mathcal{F}_C$. Since $f'(x), f''(x) \geqq 0$ for all $x \in G$ so [B] holds for f', f''. As $f = f' - f''$ it follows that [B] holds for f, too, with $O_1(f, \varepsilon) = O_1(f', \frac{\varepsilon}{2}) \cdot O_1(f'', \frac{\varepsilon}{2})$.

[C] The ξ_0 of [B] is unique. We denote it by $\overline{\mathbb{M}}(f)$.

Indeed: Let $f \in \mathcal{F}_C$ and ξ_0', ξ_0'' both fulfill [B] say with $O_1'(f, \varepsilon), O_1''(f, \varepsilon)$ respectively. Form $O = O_1'(f, \frac{\varepsilon}{2}) \cdot O_1''(f, \frac{\varepsilon}{2})$ and choose an $\mathbb{M}(g)$ in $\mathcal{M}$. Then application of [B] with ξ_0' and with ξ_0'' results in $|\xi_0' - \xi_0''| \leqq \varepsilon$. Since this is true for all $\varepsilon > 0$, therefore $\xi_0' = \xi_0''$.

[D] The $\overline{\mathbb{M}}(f)$ of [C] is a mean (in C).

Indeed: [B] and the uniqueness statement of [C] permit us to verify the requirements (α)–(γ) of the Definition in §9 immediately.

[E] The mean $\overline{\mathbb{M}}(f)$ fulfills our requirement (γ). For an $f \in \overline{\mathcal{F}}_C$ we can choose $O_1(f, \varepsilon) = O_0(f, \varepsilon)$.

Indeed: This is immediate by [B] and by [A].

[F] The mean $\overline{\mathbb{M}}(f)$ is the only mean which fulfills our requirement (γ).

Indeed: This is immediate by [C] since (γ) coincides with [B].

[G] The mean $\overline{\mathbb{M}}(f)$ is l.i.

Indeed: Consider a P with $1 \in P$. Consider an $a \in G$ and an $f(x)$ for which $f(x)$ and $f_a(x) = f(ax)$ both belong to $\overline{\mathcal{F}}_C$. For any $\varepsilon > 0$ form $O = P \cdot O_1(f, \frac{\varepsilon}{2}) \cdot O_1(f_a, \frac{\varepsilon}{2})$ and choose an $\mathbb{M}(g)$ in $\mathcal{M}_O$. Then $\mathbb{M}(g)$ is O-s.l.i., so

$$|\mathbb{M}_x(f(ax)) - \mathbb{M}_x(f(x))| \leqq K_0 \operatorname{Osc}_O(f) \leqq K_0 \operatorname{Osc}_P(f).$$

Hence application of (γ) to f and to f_a gives

$$|\overline{\mathbb{M}}x(f(ax)) - \overline{\mathbb{M}}_x(f(x))| \leqq K_0 \operatorname{Osc}_P(f) + \varepsilon.$$

Since this is true for all $\varepsilon > 0$, therefore

$$|\overline{\mathbb{M}}_x(f(ax)) - \overline{\mathbb{M}}_x(f(x))| \leqq K_0 \operatorname{Osc}_P(f).$$

So $\overline{\mathbb{M}}(g)$ is P-a.l.i.

Since this is true for all P with $1 \in P$ therefore $\overline{\mathbb{M}}(g)$ is l.i. by the lemma in §10.

[D], [E], [F], and [G] together complete the proof.

The above lemma requires one more implementation. As it stands now it does not exclude the possibility of $\overline{\mathbb{M}}(f)$ being identically 0—i.e. its correspondent $\tau(M)$ being identically 0. If this happened, then the l.i. thus obtained would of course be worthless.

One might try to exclude this possibility by requiring the existence of a fixed $\alpha > 0$ such that

(5) $$|||M||| \geqq \alpha > 0 \quad \text{for all } \mathbb{M}(f) \text{ in all } \mathcal{M}_O.$$

It is possible, however, to construct examples which fulfill $(\alpha), (\beta)$ in the above Lemma I, as well as (5), and for which $\overline{\mathbb{M}}(f)$ is nevertheless identically 0.

Consequently we must look for a different safeguard. The lemma which follows gives an obvious one:

LEMMA II. *Assume that there exists under the hypotheses of Lemma* I *above, a (fixed)* $f_0 \in \mathcal{F}_C$ *and a (fixed)* $\beta_0 > 0$ *with the following property:*

There exist for every O *with* $1 \in O$ *an* $O' = O'(O)$ *with* $1 \in O' \leqq O$ *and an* $\mathbb{M}' = \mathbb{M}'_O \in \mathcal{M}_O$, *such that*

$$|\mathbb{M}'(f_0)| \geqq \beta_0 > 0.$$

Then we have for the $\overline{\mathbb{M}}(f)$ *of Lemma I*

$$|\overline{\mathbb{M}}(f_0)| \geqq \beta_0 > 0,$$

too, and therefore this $\overline{\mathbb{M}}(f)$ *is not identically* 0.

PROOF. For any $\varepsilon > 0$ form $O = O' \cdot O_1(f_0, \varepsilon)$. Then obviously $|\overline{\mathbb{M}}(f_0)| \geqq \beta_0 - \varepsilon$. Since this is true for all $\varepsilon > 0$, so $|\overline{\mathbb{M}}(f_0)| \geqq \beta$, whence all our assertions follow.

15. Examples of means

Consider the C of §4, assuming $C^i \neq \theta$.

Consider any finite set

(1) $$\overline{F} = \overline{F}_{(m)} = (a_1, \ldots, a_m) \leqq C.$$

Then

$$\text{(2)} \qquad \mathbb{M}(f) = \sum_{i=1}^{m} \alpha_i f(a_i) \quad \text{with all } \alpha_i \geqq 0$$

is obviously a mean. And

$$\text{(3)} \qquad \mathbb{N}(f) = \frac{1}{m}\sum_{i=1}^{m} f(a_i)$$

is clearly a special case of (2).

We establish now some properties of these means.

LEMMA I. *If the set $\overline{F}$ of* (1) *is O-equidistributed, then the mean* $\mathbb{N}(f)$ *of* (3) *is O-a.l.i.* (*We choose $K_0 = 2$.*)

PROOF. Consider an O-equidistributed $\overline{F}$ of (1), and form the mean $\mathbb{N}(f)$ of (3). We must prove that (ϑ_O^*) in Definition II in §9 is fulfilled. We will find $K_0 = 2$ (cf. loc. cit.).

Consider therefore an arbitrary a and an $f(x)$ for which $f(x), f(ax)$ both belong to $\overline{\mathcal{F}}_C$.

Apply the definition of §4 to the $a_i, \ldots, a_m$ of (1) above, which we assumed to be O-equidistributed. Let us discuss the two alternatives (A) and (B).

For (A): In this case we can choose a $z_i \in a_i \cdot O \cdot aa_{\bar{i}}O$. Then $a_i^{-1}z_i$, $a_{\bar{i}}^{-1}a^{-1}z_i \in O$; hence $|f(a_i) - f(z)|, |f(aa_{\bar{i}}) - f(z_i)| \leqq \mathrm{Osc}_O(f)$, and so

$$\text{(4)} \qquad |f(a_i) - f(aa_{\bar{i}})| \leqq 2\,\mathrm{Osc}_O(f).$$

For (B): In this case we can choose a $u_i \in a_iO(-aC)$ and a $v_i \in a_{\bar{i}}O(-a^{-1}C)$. Then $a_i^{-1}a_i, a_{\bar{i}}^{-1}a^{-1}av_i$ $(= a_{\bar{i}}^{-1}v_i) \in O$ and $u_i \notin aC$, $av_i \notin C$. Hence $|f(a_i) - f(u_i)|$, $|f(aa_{\bar{i}}) - f(av_i)| \leqq \mathrm{Osc}_O(f)$, and (as $f(ax), f(x)$ belong to $\overline{\mathcal{F}}_C$), $f(u_i), f(av_i) = 0$. Therefore (4) is again valid.

Thus (4) holds for all $i = 1, \ldots, m$. Now

$$\begin{aligned}
\mathbb{N}_x(f(ax)) - \mathbb{N}_x(f(x)) &= \frac{1}{m}\sum_{i=1}^{m} f(a_i) - \frac{1}{m}\sum_{i=1}^{m} f(aa_i) \\
&= \frac{1}{m}\sum_{i=1}^{m} f(a_i) - \frac{1}{m}\sum_{i=1}^{m} f(aa_{\bar{i}}) \\
&= \frac{1}{m}\sum_{i=1}^{m} (f(a_i) - f(aa_{\bar{i}})),
\end{aligned}$$

and consequently (4) implies

$$\text{(5)} \qquad |\mathbb{N}_x(f(ax)) - \mathbb{N}_x(f(x))| \leqq 2\,\mathrm{Osc}_O(f).$$

This is precisely what we wanted and the proof is therefore completed.

LEMMA II. (A) *There exists an $f_0 \in \overline{\mathcal{F}}_C$ which is not identically* 0.

(B) *Choose a fixed $f_0 \in \overline{\mathcal{F}}_C$ by* (A). *Then there exists an $O_0 = O_0(f_0)$ with $1 \in O_0$ and a $\gamma_0 = \gamma_0(f_0) > 0$, with the following properties*:

If $1 \in O \leqq O_0$ and if the $\overline{F}$ of (1) *is O-equidistributed, then we have for the mean* $\mathbb{N}(f)$ *of* (3)

$$\mathbb{N}(f_0) \geqq \gamma_0 > 0.$$

PROOF. Ad (A): Since $C^i \neq \theta$, choose an $a_0 \in C^i$. Then there exists an everywhere continuous function $f_0(x)$ with

(i) $f_0(a_0) = 1$,

(ii) $f_0(x) = 0$ for $x \notin C^i$,

(iii) $0 \leqq f_0(x) \leqq 1$ for all $x \in G$.

(Cf. footnote 2 on p. 104. Our $(a_0), C^i$ stand for the D, θ there.) This f_0 obviously meets all requirements.

Ad (B): Consider an f_0 which satisfies (A). Then $\|f_0\| > 0$. Let P_0 be the set of all x with $f_0(x) > \frac{1}{2}\|f_0\|$. P_0 is open, since f_0 is continuous. $P_0 \leqq C$, $P_0 \neq \theta$ are obvious.

Now apply the lemma of §4 to this P_0. Then we can see, assuming $1 \in O \leqq O_0 = O_0(P_0) = O^0(f_0)$, that

$$\begin{aligned} \mathbb{N}(f_0) &= \frac{1}{m}\sum_{i=1}^{m} f_0(a_i) \geqq \frac{1}{m}\sum_{i=1}^{m}(a_i \in P_0) f_0 a_i \\ &\geqq \frac{1}{m}\sum_{i=1}^{m}(a_i \in P_0)\frac{1}{2}\|f_0\| \\ &= \frac{1}{2}\|f_0\| \cdot \frac{\text{Number of } i = 1, \ldots, m \text{ with } a_i \in P_0}{m} \\ &\geqq \frac{1}{2}\|f_0\| \cdot \beta_0. \end{aligned}$$

So we have for $\gamma_0 = \gamma_0(f_0) = \frac{1}{2}\|f_0\|\beta_0 = \frac{1}{2}\|f_0\|\beta_0(C, P_0) > 0$ the desired inequality $\mathbb{N}(f_0) \geqq \gamma_0$.

With the help of these two lemmas we obtain now:

LEMMA III. *Consider an f_0 of* (A) *in Lemma* II, *and the corresponding O_0 and γ_0 of* (B) *in Lemma* II.

Assume $1 \in O \leqq O_0$ and that the $\overline{F}$ of (1) *is O-equidistributed. Form the mean $\mathbb{N}(f)$ of* (3), *and then put*

$$\mathbb{M}(f) = \frac{\gamma_0}{\mathbb{N}(f_0)}\mathbb{N}(f).$$

Then we have

(α) $\mathbb{M}(f)$ *is a mean of the form* (2).

(β) $\mathbb{M}(f)$ *is O-a.l.i.* (*with $K_0 = 2$, as in Lemma* I).

(γ) $|||\mathbb{M}||| \leqq 1$.

(δ) $\mathbb{M}(f_0) = \gamma_0$.

PROOF. Ad (α): Obvious since $\mathbb{N}(f)$ has the form (3).

Ad (β): $\mathbb{N}(f)$ is O-a.l.i. by Lemma I with $K_0 = 2$; hence the same is true for $\mathbb{M}(f)$, as the factor $\gamma_0/\mathbb{N}(f_0)$ is $\leqq 1$ by (B) in Lemma II.

Ad (γ): Since $\mathbb{N}(f)$ has the form (3), $|\mathbb{N}(f)| \leqq \|f\|$ ensues, i.e. $|||\mathbb{N}||| \leqq 1$. This implies $|||\mathbb{M}||| \leqq 1$, since the factor $\gamma_0/\mathbb{N}(f_0)$ is $\leqq 1$ (cf. above).

Ad (δ): Obvious, owing to the definition of $\mathbb{M}(f)$.

16. Examples of means (concluded)

We are now in a position to define the sets $\mathcal{M}_O$ $(1 \in O)$ which are the substratum of the Lemmas I, II in §14.

DEFINITION I. Form f_0, O_0, γ_0 as in Lemmas II, III in §15.

Assume first $1 \in O \leqq O_0$.

Consider all O-equidistributed sets $\overline{F}$; (cf. (1) in §15) form with their help the means $\mathbb{N}(f)$ (cf. (3) in §15) and then the means

$$\mathbb{M}(f) = \frac{\gamma_0}{\mathbb{N}(f_0)} \mathbb{N}(f).$$

(cf. Lemma III in §15). Then $\mathcal{M}'_O$ is the set of all these $\mathbb{M}(f)$.

DEFINITION II. Assume now only $1 \in O$.

Then $\mathcal{M}_O = \mathcal{M}'_{O \cdot O_0}$. (We use Definition I. Clearly $1 \in O \cdot O_0 \leqq O_0$. Also $O \cdot O_0 \leqq O$.)

Lemma. $\mathcal{M}_O$ *fulfills the condition* (α) *in Lemma* I *in* §14 *and also the extra requirements of Lemma* II *in* §14.

PROOF. Ad (α) in Lemma I in §14: $\mathcal{M}_O = \mathcal{M}'_{O \cdot O_{00}}$ is not empty since $O \cdot O_{00}$-equidistributed sets $\overline{F}$ exist: e.g. by virtue of the lemmas of §5 and of §6—or equivalently by the discussion of §7. Every $\mathbb{M}(f)$ in $\mathcal{M}_O = \mathcal{M}_{O \cdot O_{00}}$ is an $O \cdot O_{00}$-a.l.i. mean by Lemma III in §15, hence a fortiori an O-a.l.i. mean.

Ad Lemma II in §14: We have for every $\mathbb{M}(f)$ in $\mathcal{M}_O = \mathcal{M}'_{O \cdot O_{00}}$ (for every O with $1 \in O$) $\mathbb{M}(f_0) = \gamma_0 > 0$ by (δ) in Lemma III in §15.

If we compare the above lemma with the conclusions reached in §14, then we see that there is only one more step needed in order to obtain an l.i. measure (in §12) by the procedure described in §14. This step consists of establishing the condition (β) in Lemma I in §14 for our above sets $\mathcal{M}_O$. The sections 17–20 which follow will be devoted to this task.

We conclude this section by reemphasizing what we observed already in the proof of our above lemma: That the equidistributed sets $\overline{F}$, which occur in the definitions and in the lemma of this section, can be obtained with the help of the lemmas of §5 and of §6—or equivalently by the discussion of §7.

17. 2-variable means

We define

DEFINITION I. $\mathcal{F}_{CC}$ is the system of those 2-variable functions $f(x, u)$ which possess the following properties:

(α) $f(x, u)$ is defined for all x, u (in G).

(β) The values of $f(x, u)$ are real numbers.

(γ) $f(x, u)$ is a continuous (2-variable) function of x, u (in all G).

(δ) $f(x, u) = 0$ whenever $x \notin C$ or $u \notin C$.

DEFINITION II. Two means $\mathbb{M}(f), \mathbb{N}(f)$ (i.e. $\mathbb{M}_x(f(x)), \mathbb{N}_x(f(x))$, cf. the definition in §9) are *commutative*: if this is true:

For every $f(x, u)$ belonging to $\mathcal{F}_{CC}$ we have:

(α) The function of x, $\mathbb{N}_u(f(x, u))$ belongs to $\mathcal{F}_C$.

(β) The function of u, $\mathbb{M}_x(f(x, u))$ belongs to $\mathcal{F}_C$.

(γ) $\mathbb{M}_x(\mathbb{N}_u(f(x, u))) = \mathbb{N}_u(\mathbb{M}_x(f(x, u)))$.

It is not difficult to prove that any two means $\mathbb{M}(f), \mathbb{N}(f)$ are commutative. Indeed: $(\alpha), (\beta)$ in the above Definition II follow with little trouble from the equivalent of Lemma I in §8 with (the "direct product group") $G \times G$ and (the "direct product set") $C \times C$ in place of its G and C. (Then $\mathcal{F}_{CC}$ replaces $\mathcal{F}_C$, i.e., we use the "uniform continuity" of $f(x, u)$.) And (γ) in the above Definition II follows from the connection between means and measures (cf. (2) and Definition II in §11) in conjunction with Fubini's theorem. We do not propose, however, to elaborate this proof here any further because this general statement is not needed for our present purposes.

All we need is this:

Lemma. *Any two means $\mathbb{M}(f), \mathbb{N}(f)$ of the form* (2) *in* §15 *are commutative.*

PROOF. I.e. we assume

(1) $$\mathbb{M}(f) = \sum_{i=1}^{m} \alpha_i f(a_i) \quad \text{(the } \alpha_i, \beta_j \text{ and the}$$

(2) $$\mathbb{N}(f) = \sum_{j=1}^{p} \beta_j f(b_j) \quad a_i, b_j\ (\in C) \text{ are fixed).}$$

Then immediately

(3) $$\mathbb{N}_u(f(x,u)) = \sum_{j=1}^{p} \beta_j f(x, b_j),$$

(4) $$\mathbb{M}_x(f(x,u)) = \sum_{i=1}^{m} \alpha_i f(a_i, u),$$

(5) $$\begin{cases} \mathbb{M}_x(\mathbb{N}_u(f(x,u))) = \mathbb{N}_u(\mathbb{M}_x(f(x,u))) \\ \qquad\qquad\qquad\quad = \displaystyle\sum_{i=1}^{m} \sum_{j=1}^{p} \alpha_i \beta_j f(a_i, b_j), \end{cases}$$

proving $(\alpha), (\beta), (\gamma)$ in our above Definition II.

Corollary. *Any two means $\mathbb{M}(f), \mathbb{N}(f)$ belonging to any two sets $\mathcal{M}_{O'}, \mathcal{M}_{O''}$ (cf. the definitions in* §16*) are commutative.*

PROOF. This follows immediately from our above lemma if we remember (α) in Lemma III in §15.

18. Comparison of two O-a.l.i. means

Consider the C of §4 assuming $C^i \neq \theta$.

Since xy is continuous (cf. §3) we can choose a P_* with $1 \in P_*$ and $P_* P_* \leqq C$. Put $O_* = P_* P_*^{-1}$; then clearly

(1) $$1 \in O_*, \quad O_* = O_*^{-1}, \quad O_* O_* \leqq C.$$

This O_* will remain fixed throughout the discussions which follow.

Assume, furthermore, that two O, P with

(2) $$1 \in O \leqq O^{\circledast}(C, P) \leqq P \leqq O_*$$

(cf. Lemma IV in §8) are given.

The discussions which follow concern two means $\mathbb{M}(f), \mathbb{N}(f)$ about which we assume for the moment only this:

(3) $$\mathbb{M}(f), \mathbb{N}(f) \text{ are } O\text{-a.l.i. means}$$

(with $K_0 = 2$, cf. the Lemmas I and III in §15).

(4) $$|||\mathbb{M}|||\,|||\mathbb{N}||| \leqq 1,$$

(5) $\mathbb{M}(f), \mathbb{N}(f)$ are commutative with themselves and with each other.

Further restrictions will be added later. Form the set $\overline{\mathcal{F}}_{O_*}$ in the sense of the definition in §8.

Clearly (by (1))

(6) $$\overline{\mathcal{F}}_{O_*} \leqq \overline{\mathcal{F}}_C \leqq \mathcal{F}_C.$$

For every function $f(x)$ form $f^{\checkmark}(x)$ as in (2) in §8:

(7) $$f^{\checkmark}(x) = f(x^{-1}).$$

Then (since $O_*^{-1} = O_*$ by (1))

(8) $f \in \overline{\mathcal{F}}_{O^*}$ is equivalent to $f^{\checkmark} \in \overline{\mathcal{F}}_{O^*}$.

(Thus in this case (6) gives $f, f^{\checkmark} \in \overline{\mathcal{F}}_C \leqq \mathcal{F}_C$. Notice that this could not be inferred e.g. from $f \in \mathcal{F}_C$.)

We proceed now to prove eight consecutive lemmas.

LEMMA I. *Assume* (1)–(5) *above. If* $f, g \in \overline{\mathcal{F}}_{O_*}$, *then*

(α) $f(x)g(u^{-1}x)$ *belongs to* $\mathcal{F}_{CC}$.

(β) $|\mathbb{M}_u(\mathbb{M}_x(f(x)g(u^{-1}x))) - \mathbb{M}(f)\mathbb{M}(g^{\checkmark})| \leqq K_0 \operatorname{Osc}_O(g^{\checkmark})$. (*Cf. Definitions* I, II *in* §11.)

PROOF. Ad (α): The requirements (α)–(γ) of Definition I in §11 are obviously satisfied so only (δ) eod.* requires closer consideration. I.e., we must prove $f(x)g(u^{-1}x) = 0$ whenever $x \notin C$ or $u \notin C$.

Assume therefore $f(x)g(u^{-1}x) \neq 0$. Then $f(x), g(u^{-1}x) \neq 0$; hence, owing to $f, g \in \overline{\mathcal{F}}_{O_*}$, $x, u^{-1}x \in O_*$. Thus by (1) $x \in O_* \leqq C$, $u = x(u^{-1}x)^{-1} \in O_*^{-1}O_*$; that is $x, u \in C$, as desired.

Ad (β): Consider a fixed $x \in C$. Assume first $f(x) \neq 0$. Then (α) above implies that $g(u^{-1}x)$ belongs to $\mathcal{F}_C$ (as a function of u—since we can divide $f(x)g(u^{-1}x)$ by $f(x)$). I.e.:

(9) $$g^{\checkmark}(x^{-1}u) = g((x^{-1}u)^{-1}) = g(u^{-1}x) \quad \text{belongs to } \mathcal{F}_C \text{ (as a function of } u).$$

Next $g^{\checkmark}(u)$ belongs to $\overline{\mathcal{F}}_{O_*}$ by (8); hence by (6)

(10) $$g^{\checkmark}(u) = g(u^{-1}) \quad \text{belongs to } \mathcal{F}_C \text{ (as a function of } u).$$

Combining (9), (10) and using (δ) in Lemma I in §9 and (ϑ_O^*) in Definition II in §10, we obtain

$$\mathbb{M}_u(f(x)g(u^{-1}x)) = f(x)\mathbb{M}_u(g(u^{-1}x)),$$

$$|\mathbb{M}_u(g(u^{-1}x)) - \mathbb{M}(g^{\checkmark})| = |\mathbb{M}_u(g^{\checkmark}(x^{-1}u)) - \mathbb{M}_u(g^{\checkmark}(u))| \leqq K_0 \operatorname{Osc}_O(g^{\checkmark});$$

hence (as $0 \leqq f(x) \leqq 1$),

(11) $$|\mathbb{M}_u(f(x)g(u^{-1}x)) - f(x)\mathbb{M}(g^{\checkmark})| \leqq K_0 \operatorname{Osc}_O(g^{\checkmark}).$$

*Editor's note: This has been replaced by eodem (in that place) in current usage.

We had to assume $f(x) \neq 0$ (for (9), hence for (11)). But if $f(x) = 0$, then both terms in the $|\cdots - \cdots|$ on the left-hand side of (11) vanish; therefore (11) is true in that case too. Thus (11) holds for all $x \in C$.

Now application of $|||\mathbb{M}||| \leqq 1$ (by (4)) gives

$$|\mathbb{M}_x(\mathbb{M}_u(f(x)g(u^{-1}x))) - f(x)\mathbb{M}(g^{\checkmark})| \leqq K_0 \operatorname{Osc}_O(g^{\checkmark});$$

i.e. (by (β) in the definition in §9 and by (δ) in Lemma I in §9—in the future we will omit such references to the elementary properties of the mean)

$$|\mathbb{M}_x(\mathbb{M}_u(f(x)g(u^{-1}x))) - \mathbb{M}(f)\mathbb{M}(g^{\checkmark})| \leqq K_0 \operatorname{Osc}_O(g^{\checkmark}). \tag{12}$$

Since $\mathbb{M}$ is commutative with itself (by (5)),

$$\mathbb{M}_u(\mathbb{M}_x(f(x)g(u^{-1}x))) = \mathbb{M}_x(\mathbb{M}_u(f(x)g(u^{-1}x))); \tag{13}$$

consequently (12) becomes

$$|\mathbb{M}_u(\mathbb{M}_x(f(x)g(u^{-1}x))) - \mathbb{M}(f)\mathbb{M}(g^{\checkmark})| \leqq K_0 \operatorname{Osc}_O(g^{\checkmark}), \tag{14}$$

as desired.

LEMMA II. *Under the hypotheses of Lemma* I *above, also*

$$\begin{aligned}\mathbb{M}_u(\mathbb{M}_x(f(x)g(u^{-1}x))) \geqq {} & \{\mathbb{M}(f)\}^2\mathbb{M}(g^{\checkmark}) \\ & - K_0(\operatorname{Osc}_O(f) + \operatorname{Osc}_O(g) + 2\operatorname{Osc}_O(g^{\checkmark})).\end{aligned}$$

PROOF. For a fixed u, $f(x)g(u^{-1}x)$ belongs to $\overline{\mathcal{F}}_C$ as a function of x because $f(x)$ does (cf. the hypothesis of Lemma I and (6)) and because $0 \leqq g(u^{-1}x) \leqq 1$ (cf. the hypothesis of Lemma I). Furthermore, $f(ux)g(x)$ belongs to $\overline{\mathcal{F}}_C$ as a function of x, because $g(x)$ does (cf. the hypothesis of Lemma I and (6)) and because $0 \leqq f(ux) \leqq 1$ (cf. the hypothesis of Lemma I). So we see:

$$f(x)g(u^{-1}x), f(ux)\cdot g(x) \quad \text{belong to } \mathcal{F}_C \text{ as functions of } x. \tag{15}$$

Besides

$$f(ux)g(u^{-1}(ux)) = f(ux)g(x). \tag{16}$$

Combining (15), (16) and using (ϑ_0^*) in Definition II in §10 we obtain

$$\begin{cases}|\mathbb{M}_x(f(ux)gx) - \mathbb{M}_x(f(x)g(u^{-1}x))| \\ \qquad \leqq K_0 \operatorname{Osc}_O(f(ux)g(x)).\end{cases} \tag{17}$$

Now clearly $\operatorname{Osc}_O(f(ux)) = \operatorname{Osc}_O(f(x))$; hence (9) in §8 gives

$$\operatorname{Osc}_O(f(ux)g(x)) \leqq \operatorname{Osc}_O(f) + \operatorname{Osc}_O(g).$$

So (17) becomes

$$\begin{cases}|\mathbb{M}_x(f(ux)g(x)) - \mathbb{M}_x(f(x)g(u^{-1}x))| \\ \qquad \leqq K_0(\operatorname{Osc}_O(f) + \operatorname{Osc}_O(g)).\end{cases} \tag{18}$$

Next observe that $0 \leqq f(ux)g(x) \leqq g(x)$ (again owing to the hypothesis of Lemma I); hence

$$\mathbb{M}_x(f(ux)g(x)) \leqq \mathbb{M}(g). \tag{19}$$

Consequently (18), (19) give

$$\mathbb{M}_x(f(x)g(u^{-1}x)) \leqq \mathbb{M}(g) + K_0(\operatorname{Osc}_O(f) + \operatorname{Osc}_O(g)).$$

Now apply $\mathbb{M}_u$ (using $|||\mathbb{M}||| \leqq 1$, cf. (4)); then this becomes

$$\mathbb{M}_u(\mathbb{M}_x(f(x)g(u^{-1}x))) \leqq \mathbb{M}(g) + K_0(\mathrm{Osc}_O(f) + \mathrm{Osc}_O(g)). \tag{20}$$

Combination of (20) with (β) in Lemma I above gives

$$\mathbb{M}(g) \geqq \mathbb{M}(f)\mathbb{M}(g^{\checkmark}) - K_0(\mathrm{Osc}_O(f) + \mathrm{Osc}_O(g) + \mathrm{Osc}_O(g^{\checkmark})). \tag{21}$$

Replace g by $g^{\checkmark}$ in (21); then this becomes (remember (8))

$$\mathbb{M}(g^{\checkmark}) \geqq \mathbb{M}(f)\mathbb{M}(g) - K_0(\mathrm{Osc}_O(f) + \mathrm{Osc}_O(g) + \mathrm{Osc}_O(g^{\checkmark})). \tag{22}$$

Now substitute (22) into (β) in Lemma I above. Observe that $\mathbb{M}(f) \leqq 1$ owing to $0 \leqq f(x) \leqq 1$ (use again $|||\mathbb{M}||| \leqq 1$, cf. (4)). Then we obtain

(23)

$$\mathbb{M}_u(\mathbb{M}_x(f(x)g(u^{-1}x))) \leqq \{\mathbb{M}(f)\}^2\mathbb{M}(g) - K_0(\mathrm{Osc}_O(f) + \mathrm{Osc}_O(g) + 2\,\mathrm{Osc}_O(g^{\checkmark}))$$

as desired.

LEMMA III. *Under the hypotheses of Lemma* I *above, there exists a function* $\varphi^+(x)$ *and* u^+ $(\in G)$ *with the following properties*:

(α) $f(x) = \varphi(x)+f^+(x)$, $g(x) = \varphi(u^+x)+g^+(x)$, *all four functions* $\varphi(x), f^+(x)$, $\varphi(u^+x), g^+(x)$ *belonging to* $\overline{\mathcal{F}}_{O_*}$.

(β) *Put* $\delta = \mathrm{Max}(\mathrm{Osc}_P(f), \mathrm{Osc}_P(g))$, $\delta^+ = \mathrm{Max}(\mathrm{Osc}_P(f^+), \mathrm{Osc}_P(g^+))$. *Then* $\delta^+ \leqq 2\delta$, $\mathrm{Osc}_P(\varphi) \leqq 2\delta$.

(γ) *Put* $\alpha = \mathbb{M}(f)\mathbb{M}(g)$, $\alpha^+ = \mathbb{M}(f^+)\mathbb{M}(g^+)$. *Then* $\alpha^+ \leqq \alpha - \alpha^2 + 4K_0\delta$.

PROOF. Form α, δ (which depend on f, g only) as indicated in $(\beta), (\delta)$ above. Then (2) gives directly

$$\mathrm{Osc}_O(f) \leqq \mathrm{Osc}_P(f), \qquad \mathrm{Osc}_O(g) \leqq \mathrm{Osc}_P(g),$$

and with the help of Lemma IV in §8

$$\mathrm{Osc}_O(g^{\checkmark}) \leqq \mathrm{Osc}_P(g).$$

Consequently

$$\mathrm{Osc}_O(f) + \mathrm{Osc}_O(g) + 2\,\mathrm{Osc}_O(g^{\checkmark}) \leqq 4\delta.$$

Hence Lemma II above gives

$$\mathbb{M}_u(\mathbb{M}_x(f(x)g(u^{-1}x)) \geqq \{\mathbb{M}(f)\}^2\mathbb{M}(g) - 4K_0\delta;$$

i.e.

$$\begin{cases} \mathbb{M}(h) \geqq \{\mathbb{M}(f)\}^2\mathbb{M}(g) - 4K_0\delta \\ \qquad \text{with } h(u) = \mathbb{M}_x(f(x)g(u^{-1}x)). \end{cases} \tag{24}$$

Thus (using $|||\mathbb{M}||| \leqq 1$, by (4))

$$\|h\| \geqq \{\mathbb{M}(f)\}^2\mathbb{M}(g) - 4K_0\delta;$$

hence there exists by Lemma III in §8 a $u^+ \in G$ with

$$h(u^+) \geqq \{\mathbb{M}(f)\}^2\mathbb{M}(g) - 4K_0\delta.$$

I.e.

$$\mathbb{M}_x(f(x)g(u^{-1}x)) \geqq \{\mathbb{M}(f)\}^2\mathbb{M}(g) - 4K_0\delta. \tag{25}$$

We now put

$$\varphi(x) = f(x)g(u^{+-1}x). \tag{26}$$

Then

(27) $$\varphi(u^+x) = f(u^+x)g(x),$$

and

(28) $$f^+(x) = f(x) - \varphi(x) = f(x)(1 - g(u^{+-1}x)),$$

(29) $$g^+(x) = g(x) - \varphi(u^+x) = g(x)(1 - f(u^+x)).$$

As $f, g \in \overline{\mathcal{F}}_{O_*}$, hence always $0 \leqq f(x), g(x) \leqq 1$, therefore (26)–(29) imply

(30) $$0 \leqq \varphi(x), \quad f^+(x), \varphi(u^+x), g^+(x) \leqq 1 \quad \text{for all } x \in G.$$

And since $x \in O_*$ implies $f(x) = g(x) = 0$, therefore (26)–(29) imply

(31) $$\varphi(x) = f^+(x) = \varphi(u^+x) = g^+(x) = 0 \quad \text{if } x \in O_*.$$

(30), (31) mean that

(32) $$\varphi(x), f^+(x), \varphi(u^+x), g^+(x) \quad \text{belong to } \overline{\mathcal{F}}_{O_*}.$$

Now we can prove (α)–(δ):

Ad (α): Obvious by (28), (29), (32).

Ad (β): Immediate by (28), (29), remembering (9) in §8.

Ad (γ): Immediate by (26), remembering (9) in §8.

Ad (δ): Always $\varphi(u^+x) \geqq 0$; therefore $\mathbb{M}_x(\varphi(u^+x)) \geqq 0$, and hence (by (26) or by (α))

(33) $$\mathbb{M}(g^+) \leqq \mathbb{M}(g).$$

(25) and (26) give (by (26) or by (α))

(34) $$\mathbb{M}(f^+) \leqq M(f) - \{\mathbb{M}(f)\}^2 \cdot \mathbb{M}(g) + 4K_0\delta.$$

Now multiply (33) and (34). Observe that $\mathbb{M}(g) \leqq 1$ owing to $0 \leqq g(x) \leqq 1$ (use again $|||\mathbb{M}||| \leqq 1$, cf. (4)). Then we obtain

(35) $$\mathbb{M}(f^+)\mathbb{M}(g^+) \leqq \mathbb{M}(f)\mathbb{M}(g) - \{\mathbb{M}(f)M(g)\}^2 + 4K_0\delta;$$

i.e.

(36) $$\alpha^+ \leqq \alpha - \alpha^2 + 4K_0\delta$$

as desired.

LEMMA IV. *Under the hypotheses of Lemma* I *above there exist a sequence of functions $\varphi_\ell(x)$ and a sequence $u_\ell \in G$, $\ell = 1, 2, \ldots$ with the following properties:*

(α)

$$\left.\begin{aligned} f(x) = f_0(x) &= \sum_{n=1}^{\ell} \varphi_n(x) + f_\ell(x) \\ g(x) = g_0(x) &= \sum_{n=1}^{\ell} \varphi_n(u_n x) + g_\ell(x) \end{aligned}\right\}, \qquad \ell = 1, 2, \ldots,$$

all four functions $\varphi_\ell(x), f_\ell(x), \varphi_\ell(u_\ell x), g_\ell(x)$ belonging to $\overline{\mathcal{F}}_{O_*}$.

(β) *Put*

$$\delta_\ell = \mathrm{Max}(\mathrm{Osc}_P(f), \mathrm{Osc}_P(g)), \qquad \ell = 1, 2, \ldots.$$

Then

$$\delta_\ell \leqq 2\delta_{\ell-1} \qquad (\ell = 1, 2, \ldots).$$

(γ)

$$\mathrm{Osc}_P(\varphi_\ell) \leqq 2\delta_{\ell-1} \qquad (\ell = 1, 2, \dots).$$

(δ) *Put*

$$\alpha_\ell = \mathbb{M}(f_\ell)\mathbb{M}(g_\ell) \qquad (\ell = 0, 1, 2, \dots).$$

Then

$$\alpha_\ell \leqq \alpha_{\ell-1} - \alpha_{\ell-1}^2 + 4K_0\delta_{\ell-1} \qquad (\ell = 0, 1, 2, \dots).$$

PROOF. Put (as indicated in (α) above)

$$f_0(x) = f(x), \qquad g_0(x) = g(x).$$

Consider an $\ell = 1, 2, \dots$ such that $f_{\ell-1}(x), g_{\ell-1}(x)$ are already defined and belonging to $\overline{\mathcal{F}}_{O_*}$.

Then apply Lemma III above with these $f_{\ell-1}(x), g_{\ell-1}(x)$ in place of its $f(x), g(x)$ and choose our $\varphi_\ell(x), u_\ell, f_\ell(x), g_\ell(x)$ as its $\varphi(x), u^+, f^+(x), g^+(x)$. Then our ($\alpha$)–($\delta$) coincide with the ($\alpha$)–($\delta$) of Lemma III above respectively. This completes the proof.

LEMMA V. *Under the hypotheses and with the notations of Lemma* IV *we have: If*

(α)

$$\delta = \delta_0 = \mathrm{Max}(\mathrm{Osc}_P(f), \mathrm{Osc}_P(g)) \leqq \frac{1}{2^{p+1}p^2(p+1)K_0} \quad \text{for a fixed } p = 2, 3, \dots,$$

then

(β)

$$\delta_\ell \leqq \frac{1}{2^{p-\ell+1}p^2(p+1)K_0} \quad \text{for } \ell = 0, 1, 2, \dots,$$

(γ)

$$\alpha_\ell \leqq \frac{1}{\ell+1} \quad \text{for } \ell = 0, 1, 2, \dots, p.$$

PROOF. Ad (β): This is an immediate consequence of (α) above, using (β) in Lemma IV.

Ad (γ): Consider first $\ell - \sigma$. Clearly $\mathbb{M}(f), \mathbb{M}(g) \leqq 1$ (we argued this repeatedly before); hence $\alpha_0 \leqq 1$, so that (γ) holds in this case.

Consider next $\ell = 1$. By (δ) in Lemma IV,

$$\alpha_1 \leqq \alpha_0 - \alpha_0^2 + 4K_0\delta_0 \leqq \frac{1}{4} + \frac{1}{2^{p-1}p^2(p+1)} \leqq \frac{1}{4} + \frac{1}{24} < \frac{1}{2},$$

so that (γ) holds in this case, too.

Consider finally an $\ell = 2, 3, \dots, p$, such that (γ) is already known to hold for $\ell - 1$. Then $\alpha_{\ell-1} \leqq 1/\ell$ and so by (δ) in Lemma IV and by (β) above

$$\begin{aligned}\alpha_\ell &\leqq \alpha_{\ell-1} - \alpha_{\ell-1}^2 + 4K_0\delta_{\ell-1} \leqq \frac{1}{\ell} - \frac{1}{\ell^2} + \frac{1}{2^{p-\ell}p^2(\ell+1)} \\ &\leqq \frac{1}{\ell} - \frac{1}{\ell} + \frac{1}{\ell^2(\ell+1)} = \frac{1}{\ell} - \frac{1}{\ell(\ell+1)} = \frac{1}{\ell+1}\end{aligned}$$

so that (γ) holds for ℓ, too.

Thus (γ) holds for all $e = 0, 1, 2, \dots, p$ and the proof is completed.

LEMMA VI. *Under the hypotheses and with the notations of Lemmas* IV, V *we have*:

There exists an $\ell = 0, 1, 2, \dots,$ *such that*

(α) $K_0 \sum_{n=1}^{\ell} \mathrm{Osc}_P(\varphi_n) \leqq \omega(\delta)$,

(β) *either* $\mathbb{M}(f_\ell) \leqq \omega(\delta)$ *or* $\mathbb{M}(g_\ell) \leqq \omega(\delta)$.

Here $\omega(\delta)$ *is a fixed numerical function with the following properties*:

(γ) $\omega(\delta)$ *is defined for the* $\delta > 0$, *and there* $0 \leqq \omega(\delta) \leqq 1$,

(δ) $\mathrm{Lim}_{\delta \to 0}\, \omega(\delta) = 0$.

REMARK. To be precise, $\omega(\delta)$ is the following function:
Put

$$\varepsilon_p = \frac{1}{2^{p+1}p^2(p+1)K_0} \quad \text{for } p = 2, 3, \dots . \tag{37}$$

Clearly

$$\frac{1}{96K_0} = \varepsilon_2 > \varepsilon_3 > \varepsilon_4 > \cdots > 0. \tag{38}$$

Now define

$$\omega(\delta) = \begin{cases} 1 & \text{for } \delta > \varepsilon_2, \\ \frac{1}{\sqrt{p}} & \text{for } \varepsilon_{p+1} < \delta < \varepsilon_p \text{ with } p = 2, 3, \dots . \end{cases} \tag{39}$$

Thus we have asymptotically

$$\omega(\delta) \sim \frac{1}{\sqrt{\log_2 1/\delta}} \quad \text{for } \delta \to 0. \tag{40}$$

PROOF. For $\delta > \varepsilon_2$: Put $\ell = 0$. Then the left-hand side of (α) vanished; hence (α) is true. In (β) remember that $\mathbb{M}(f), M(g) \leqq 1$; hence ($\beta$) is also true.

For $\varepsilon_{p+1} < \delta \leqq \varepsilon_p$ with $p = 2, 3, \dots$: Put $\ell = p$. Then Lemmas IV, V give

$$\begin{aligned} K_0 \sum_{h=1}^{p} \mathrm{Osc}_P(\varphi_h) \leqq K_0 \sum_{h=1}^{p} 2\delta_{h-1} &\leqq \frac{1}{2^{p-h+1}p^2(p+1)} \\ &= \frac{1}{p^2(p+1)} \sum_{h=1}^{p} \frac{1}{2^{p-h+1}} < \frac{1}{p^2(p+1)} < \frac{1}{\sqrt{p}}, \end{aligned}$$

$$K_0 \sum_{h=1}^{p} \mathrm{Osc}_p(\varphi_h) < \frac{1}{\sqrt{p}}. \tag{41}$$

And $\alpha_p \leqq \frac{1}{p+1} < \frac{1}{p}$, $\mathbb{M}(f_p)\mathbb{M}(g_p) < \frac{1}{p}$; hence

$$\begin{cases} \text{either} & \mathbb{M}(fp) < \frac{1}{\sqrt{p}} \\ \text{or} & \mathbb{M}(gp) < \frac{1}{\sqrt{p}}. \end{cases} \tag{42}$$

Now (41), (42) prove (α), (β) respectively.
Thus the proof is completed.

LEMMA VII. *Assume, as before* (1)–(5). *We will, however, use now both means* $\mathbb{M}(f), \mathbb{N}(f)$ *of* (3)–(5). *If* $f, g \in \overline{\mathcal{F}}_{O_*}$, *then either*

(A) *both* $\mathbb{M}(f) \geqq \mathbb{M}(g) - 3\omega(\delta)$, $\mathbb{N}(f) \geqq \mathbb{N}(g) - 3\omega(\delta)$ *or*

(B) *both* $\mathbb{M}(f) \leqq \mathbb{M}(g) + 3\omega(\delta)$, $\mathbb{N}(f) \leqq \mathbb{N}(g) + 3\omega(\delta)$.

($\omega(\delta)$ *is the function introduced in Lemma* VI, *in particular in the remark following that lemma.*)

PROOF. Form the mean

$$\mathbb{M}'(f) = \frac{1}{2}(\mathbb{M}(f) + \mathbb{N}(f)).$$

It is clear that $\mathbb{M}'(f)$ fulfills (3)–(5) along with $\mathbb{M}(f), \mathbb{N}(f)$.

Now apply Lemma VI, with $\mathbb{M}'(f)$ in place of its $\mathbb{M}(f)$. Choose the $\ell = 0, 1, 2, \ldots$ as indicated there; then $(\alpha), (\beta)$ of Lemma Vi hold. We may assume by symmetry (for f, g), that the second alternative of (β) holds. From this we are going to derive (A) above.

(α) states:

$$K_0 \sum_{h=1}^{\ell} \mathrm{Osc}_P(\varphi_h) \leqq \omega(\delta). \tag{43}$$

The second alternative states with $\mathbb{M}'(f) = \frac{1}{2}(\mathbb{M}(f) + \mathbb{N}(f))$ in place of its $\mathbb{M}(f)$

$$\frac{1}{2}(\mathbb{M}(g_\ell(x)) + \mathbb{N}(g_\ell(x))) \leqq \omega(\delta).$$

This implies, since $\mathbb{M}(g_\ell(x)), \mathbb{N}(g_\ell(x)) \geqq 0$

$$\mathbb{M}(g_\ell(x)), \mathbb{N}(g_\ell(x)) \leqq 2\omega(\delta). \tag{44}$$

Add to these the obvious inequalities

$$\mathbb{M}(f_\ell(x)), \mathbb{N}(f_\ell(x)) \geqq 0. \tag{45}$$

Finally, (ϑ_0^*) in Definition II in §10 gives

$$\begin{aligned} \mathbb{M}(\varphi_h(u_h x)) &\leqq \mathbb{M}(\varphi_h(x)) + K_0 \,\mathrm{Osc}_O(\varphi_h) \\ &\leqq \mathbb{M}(\varphi_h(x)) + K_0 \,\mathrm{Osc}_P(\varphi_h); \end{aligned}$$

hence

$$\sum_{h=1}^{\ell} \mathbb{M}(\varphi_h(u_h x)) \leqq \sum_{h=1}^{\ell} \mathbb{M}(\varphi_h(x)) + K_0 \sum_{h=1}^{\ell} \mathrm{Osc}_p(\varphi_h). \tag{46}$$

Similarly

$$\sum_{h=1}^{\ell} \mathbb{N}(\varphi_h(u_h x)) \leqq \sum_{h=1}^{\ell} \mathbb{N}(\varphi_h(x)) + K_0 \sum_{h=1}^{\ell} \mathrm{Osc}_P(\varphi_h). \tag{47}$$

Now a combination of (α) in Lemma IV with (44), (45) and (46), (47) and (43) gives

$$\begin{cases} \mathbb{M}(g) \leqq \mathbb{M}(f) + 3\omega(\delta), \\ \mathbb{N}(g) \leqq \mathbb{N}(f) + 3\omega(\delta), \end{cases} \tag{48}$$

which coincides with the two inequalities of (A) as desired.

LEMMA VIII. *Under the hypotheses and with the notations of Lemma* VII *we have*:

$$|\mathbb{M}(f)\mathbb{N}(g) - \mathbb{M}(g)\mathbb{N}(f)| \leqq 6\omega(\delta).$$

PROOF. Consider a real constant γ_0 with

$$0 \leqq \gamma_0 \leqq 1. \tag{49}$$

We may apply Lemma VII to $\gamma_0\mathbb{M}(f), \mathbb{N}(f)$ in place of its $\mathbb{M}(f), \mathbb{N}(f)$ since the former obviously satisfy (3)–(5) along with the latter. Hence we cannot have

$$\gamma_0\mathbb{M}(f) < \mathbb{M}(g) - 3\omega(\delta), \tag{50}$$

$$\gamma_0\mathbb{N}(f) > \mathbb{N}(g) + 3\omega(\delta), \tag{51}$$

simultaneously, since (50), (51) would exclude both alternatives (A), (B) of Lemma VII.

In other words, $\mathbb{N}_0\gamma_0$ can fulfill all three conditions (49), (50), (51) at the same time.

It is clear, however, that a γ_0 with

$$\frac{\mathbb{N}(g) + 3\omega(\delta)}{\mathbb{N}(f)} < \gamma_0 \begin{cases} < \dfrac{\mathbb{M}(g) - 3\omega(\delta)}{\mathbb{M}(f)} \\ \leqq 1 \end{cases}$$

would fulfill (50), (51). And such a γ_0 would exist if

$$\frac{\mathbb{N}(g) + 3\omega(\delta)}{\mathbb{N}(f)} \begin{cases} < \dfrac{\mathbb{M}(g) - 3\omega(\delta)}{\mathbb{M}(f)} \\ \leqq 1. \end{cases} \tag{52}$$

Hence (52) must not be true.

If the first inequality of (52) is not true, then we have

$$\mathbb{M}(f)(\mathbb{N}(g) + 3\omega(\delta)) \geqq \mathbb{N}(f)(\mathbb{M}(g) - 3\omega(\delta)),$$
$$\mathbb{M}(f)\mathbb{N}(g) - \mathbb{M}(g)\mathbb{N}(f) \geqq 3(\mathbb{M}(f) + \mathbb{N}(f))\omega(\delta).$$

Since $\mathbb{M}(f), N(f) \leqq 1$ (we argued this repeatedly before) the above inequality implies

$$\mathbb{M}(f)\mathbb{N}(g) - \mathbb{M}(g)\mathbb{N}(f) \geqq -6\omega(\delta). \tag{53}$$

If the second inequality of (52) is not true, then we have

$$\mathbb{N}(f) < \mathbb{N}(g) + 3\omega(\delta). \tag{54}$$

Hence we see: Either (53) or (54) is true.

Assume now that (53) is not true. Then (54) holds. A simultaneous interchange of $\mathbb{M}, \mathbb{N}$ and of f, g does not affect our original hypotheses, nor (53). Hence its application to (54) is also true; i.e.

$$\mathbb{M}(g) < \mathbb{M}(f) + 3\omega\delta. \tag{55}$$

Now (54), (55) give, remembering $\mathbb{M}(f), \mathbb{N}(f) \leqq 1$ (cf. above),

$$\begin{aligned} \mathbb{M}(f)\mathbb{N}(g) - \mathbb{M}(g)\mathbb{N}(f) &= \mathbb{M}(f)\{\mathbb{N}(g) - \mathbb{N}(f)\} + \mathbb{N}(f)\{\mathbb{M}(f) - \mathbb{M}(g)\} \\ &> 1 \cdot \{-3\omega(\delta)\} + 1 \cdot \{-3\omega(\delta)\} \\ &= -6\omega(\delta), \end{aligned}$$

i.e. (53).

Thus (53) is implied by its own negation, I.e. (53) is true unconditionally.

Interchanging $\mathbb{M}, \mathbb{N}$ (but not f, g) does not affect our original hypotheses. Hence its application to (53) is also true, i.e.

$$\mathbb{M}(f)\mathbb{N}(g) - \mathbb{M}(g)\mathbb{N}(f) \leqq 6\omega(\delta). \tag{56}$$

(53), (56) together give the desired inequality and thus complete the proof.

19. Comparison of two O-a.l.i. means (concluded)

We continue all assumptions formulated at the beginning of §18 (before its Lemma I). We propose to extend Lemma VIII in §18 from the $f, g \in \overline{\mathcal{F}}_{O_*}$ to which it was restricted (cf. the Lemmas VII, VIII in §18) to all $f, g \in \overline{\mathcal{F}}_C$ in a somewhat modified form.

For this purpose we need a preparatory lemma.

LEMMA I. *Given C and a $Q \neq \theta$ there exist a $t = t(C,Q)$ $(= 1, 2, \ldots)$ and t functions $f_s(x) = f_s(C, Q; x)$, $s = 1, \ldots, t$ and t corresponding elements $a_s = a_s(C, Q)$ $(\in G)$, $s = 1, \ldots, t$ with the following properties*:

(α) *$f_s \in \overline{\mathcal{F}}_Q$ for all $s = 1, \ldots, t$.*

(β) *$\sum_{s=1}^{t} f_s(a_s x) = 1$ if $x \in C$.*

PROOF. Consider an $x_0 \in Q$. C is regular and locally compact, so choose a Q' with $x_0 \in Q' \leqq \overline{Q'} \leqq Q$, $\overline{Q'}$ compact. Then there exists an everywhere continuous function $h(x)$ with

(i) $h(x) = 1$ for $x \in \overline{Q'}$,

(ii) $h(x) = 0$ for $x \notin Q$,

(iii) $0 \leqq h(x) \leqq 1$ for all $x \in G$.

(Cf. footnote 2 on p. 104. Our $\overline{Q'}, Q$ stand for the D, O there.)

We have $C \leqq \sum_{s=1}^{t} b_s Q'$. (This is due to $C \leqq \sum_a aQ'$, cf. the arguments made in §8.) Consider the function

$$k(x) = \sum_{s=1}^{t} h(b_s^{-1} x). \tag{1}$$

$k(x)$ is everywhere continuous since $h(x)$ is. If $x \in C$ then $x \in b_{s_1} Q'$ for some $s_1 = 1, \ldots, t$. Hence $b_{s_1}^{-1} x \in Q'$, $h(b_{s_1}^{-1} x) = 1$, and consequently $k(x) \geqq h(b_{s_1}^{-1} x) = 1$. Thus

$$k(x) \geqq 1 \quad \text{for all } x \in C. \tag{2}$$

Now put

$$f_s(x) = \frac{h(x)}{\operatorname{Max}(1, k(b_s x))}. \tag{3}$$

Our (α), (β) holds for this t and these f_s, $s = 1, \ldots, t$, if we put $a_s = b_s^{-1}$, $s = 1, \ldots, t$. Indeed:

Ad (α): $h \in \overline{\mathcal{F}}_Q$ by (i)–(iii) and it implies $b_s \in \overline{\mathcal{F}}_Q$, and as the denominator in (3) is everywhere continuous and $\geqq 1$.

Ad (β): For $x \in C$ we have, using (2), (3):

$$\sum_{s=1}^{t} f_s(b_s^{-1} x) = \sum_{s=1}^{t} \frac{h(b_s^{-1} x)}{\operatorname{Max}(1, k(x))} = \sum_{s=1}^{t} \frac{h(b_s^{-1} x)}{k(x)} = 1$$

as desired.

Now the extension of Lemma VIII in §18 which we envisaged can be carried out.

LEMMA II. *Assume, as before* (1)–(5) *in* §18, *using both means* $\mathbb{M}(f), \mathbb{N}(f)$ *of* (3)–(5) *in* §18. *Then there exist a positive, finite constant* $\mathcal{C} = \mathcal{C}(C)$, *a* $t = t(c)$ $(= 1, 2, \ldots,)$, *and* t *functions* $f_s(x) = f_s(C; x)$, $s = 1, \ldots, t$, *with the following properties*:

(α) $f_s \in \overline{\mathcal{F}}_C$ *for all* $s = 1, 2, \ldots, t$.

(β) *If* $f, g \in \overline{\mathcal{F}}_C$, *then*

$$|\mathbb{M}(f)\mathbb{N}(g) - \mathbb{M}(g)\mathbb{N}(f)| \leqq \mathcal{C}\omega(2\eta),$$

where

$$\eta = \operatorname{Max}(\operatorname{Osc}_P(f), \operatorname{Osc}_P(g), \operatorname{Osc}_P(f_1), \ldots \operatorname{Osc}_P(f_t)).$$

($\omega(\delta)$ *is the function introduced in Lemma* VI *in* §18, *in particular in the remark following that lemma.*)

PROOF. Apply Lemma I above to C and to $Q = O_*$ (O_* from (1) in §18), and form the t, and the f_s, $s = 1, \ldots, t$, accordingly. In Lemma I these depended on C, Q but now they depend on C only, since C determines the $Q = O_*$.

We define next

$$\left.\begin{aligned} f_s'(x) &= f(a_s^{-1}x) f_s(x) \\ g_s'(x) &= g(a_s^{-1}x) f_s(x) \end{aligned}\right\} \quad \text{for } s = 1, \ldots, t. \tag{4}$$

Using Lemma I we see that $f_s'(x), g_s'(x)$ belong to $\overline{\mathcal{F}}_{O_x}$ along with $f_s(x)$; i.e.

$$f_s', g_s' \in \overline{\mathcal{F}}_{O_*} \leqq \overline{\mathcal{F}}_C \text{ for } s = 1, \ldots, t. \tag{5}$$

And since

$$\left.\begin{aligned} f_s'(a_s x) &= f(x) f_s(a_s x) \\ g_s'(a_s x) &= g(x) f_s(a_s x) \end{aligned}\right\} \quad \text{for } s = 1, \ldots, t; \tag{6}$$

therefore $f_s'(a_s x), g_s'(a_s x)$ belong to $\overline{\mathcal{F}}_C$ along with $f(x), g(x)$; i.e.

$$f_s'(a_s x), g_s'(a_s x) \quad \text{belong to } \overline{\mathcal{F}}_C \text{ for } s = 1, \ldots, t. \tag{7}$$

Clearly $\operatorname{Osc}_P(f(a_s^{-1}x)) = \operatorname{Osc}_P(f(x))$; hence (9) in §8 gives, considering (4),

$$\operatorname{Osc}_P(f_s') \leqq \operatorname{Osc}_P(f) + \operatorname{Osc}_P(f_s).$$

Similarly

$$\operatorname{Osc}_P(g_s') \leqq \operatorname{Osc}_P(f) + \operatorname{Osc}_P(f_s).$$

Remembering the definition of η these become

$$\operatorname{Osc}_P(f_s'), \operatorname{Osc}_P(g_s') \leqq 2\eta. \tag{8}$$

Since always $0 \leqq f_s'(x), g_s'(x) \leqq 1$, by (5) and so $0 \leqq f_s'(a_s x), g_s'(a_s x) \leqq 1$, therefore $|||\mathbb{M}|||, |||\mathbb{N}||| \leqq 1$ (remember (4) in §18) gives

$$\left\{\begin{aligned} &0 \leqq \mathbb{M}_x(f_s'(x)), \mathbb{N}_x(f_s'(x)), \mathbb{M}_x(f_s'(a_s x)), \mathbb{N}_x(f_s'(a_s x)), \\ &\quad \mathbb{M}_x(g_s'(x)), \mathbb{N}_x(g_s'(x)), \mathbb{M}_x(g_s'(a_s x)), \mathbb{N}_x(g_s'(a_s x)) \leqq 1. \end{aligned}\right. \tag{9}$$

Observe finally that

$$
(10) \qquad \left.\begin{aligned} f(x) &= \sum_{s=1}^{t} f'_s(a_s x) \\ g(x) &= \sum_{s=1}^{t} g'_s(a_s x) \end{aligned}\right\} \quad \text{for all } x \in G.
$$

Indeed: For $x \in C$ this is a consequence of (6) and of (β) in Lemma I. For $x \notin C$ both sides vanish, since $f, g \in \overline{\mathcal{F}}_C$ and considering (7).

We will now show that our $(\alpha), (\beta)$ hold for this t and these f_s, $s = 1, \ldots, t$, if we put $\mathcal{C} = 14t^2$. Indeed:

Ad (α): Owing to (α) in Lemma I, $f_s \in \overline{\mathcal{F}}_{O_*} \leqq \overline{\mathcal{F}}_C$.

Ad (β): Owing to (5), (7)

$$
|\mathbb{M}_x(f'_s(a_s x)) - \mathbb{M}_x(f'_s(x))| \leqq K_0 \operatorname{Osc}_P(f'_s);
$$

hence by (8)

$$
(11) \qquad |\mathbb{M}_x(f'_s(a_s x)) - \mathbb{M}_x(f'_s(x))| \leqq 2K_0\eta.
$$

Also by (9)

$$
(12) \qquad |\mathbb{M}_x(f'_s(a_s x)) - \mathbb{M}_x(f'_s(x))| \leqq 2.
$$

Now inspection of the definitory remark after Lemma VI in §18 shows that

$$
(13) \qquad \omega(\delta) \geqq \operatorname{Min}\left(1, \frac{K_0\delta}{2}\right) \quad \text{for all } \delta > 0.
$$

Application of (13) (with $\delta = 2\eta$) to (11), (12) gives therefore

$$
(14) \qquad |\mathbb{M}_x(f'_s(a_s x)) - \mathbb{M}_x(f'_s(x))| \leqq 2\omega(2\eta).
$$

Replacement of f, f'_s by g, g'_s (the f_s and the a_s, remain unchanged!) transforms (14) into

$$
(15) \qquad |\mathbb{M}_x(g'_s(a_s x)) - \mathbb{M}_x(g'_s(x))| \leqq 2\omega(2\eta).
$$

Replacement of $\mathbb{M}$ by $\mathbb{N}$ (everything else remains unchanged!) transforms (14), (15) into

$$
(16) \qquad |\mathbb{N}_x(f'_s(a_x s)) - \mathbb{N}_x(f'_s(x))| \leqq 2\omega(2\eta),
$$

$$
(17) \qquad |\mathbb{N}_x(g'_s(a_s x)) - \mathbb{N}_x(g'_s(x))| \leqq 2\omega(2\eta).
$$

(9) and (14)–(17) give

$$
(18) \qquad \begin{aligned} &|\{\mathbb{M}_x(f'_s(a_s x))\mathbb{N}_n(g'_s(a_s x)) - \mathbb{M}_x(g'_s(a_s x))\mathbb{N}_x(f'_s(a_s x))\} \\ &\quad - \{\mathbb{M}_x(f'_s(x))\mathbb{N}_x(g'_s(x)) - \mathbb{M}_x(g'_s(x))\mathbb{N}_x(f'_s(x))\}| \leqq 8\omega(2\eta). \end{aligned}
$$

(5) and Lemma VIII in §18 give, considering (8),

$$
(19) \qquad |\mathbb{M}_x(f'_s(a_s x))\mathbb{N}_x(g'_s(a_s x)) - \mathbb{M}_x(g'_s(a_s x))\mathbb{N}_x(f'_s(a_s x))| \leqq 6\omega(2\eta).
$$

Now (18), (18) give together

$$
(20) \qquad |\mathbb{M}(f'_s)\mathbb{N}(g'_s) - \mathbb{M}(g'_s)\mathbb{N}(f'_s)| \leqq 14\omega(2\eta).
$$

And now (10) and (20)

$$|\mathbb{M}(f)\mathbb{N}(g) - \mathbb{M}(g)\mathbb{N}(f)| \leqq \sum_{s,u=1}^{t} |\mathbb{M}(f)\mathbb{N}(g'_u) - \mathbb{M}(g'_u)\mathbb{N}(f'_s)|$$
$$\leqq \sum_{s,u=1}^{t} 14\omega(2\eta) = 14t^2\omega(2\eta),$$

i.e.

$$|\mathbb{M}(f)N(g) - \mathbb{M}(g)\mathbb{N}(f)| \leqq 14t^2\omega(2\eta).$$

This is the desired statement of (β).

20. The convergence theorem

Consider the C of §4, assuming $C^i \neq \theta$.

We are now able to fill in the gap indicated in the discussion at the end of §16, by proving Lemma I.

The $\mathcal{M}_O$ of the Definitions I, II, in §16 fulfill all requirements of Lemmas I, II, in §14, as we know from lemmas of §16, so that we must only verify the condition (β) of Lemma I in §14. Certain (fixed) f_0, O_0, γ_0 are given by Definition I in §16.

Let an $f \in \overline{\mathcal{F}}_C$ and an $\varepsilon > 0$ be given.

Choose an $\eta_0 > 0$ so that (with the C, the t, and the f_s, $s = 1, \ldots, t$, of Lemma II in §19 for γ_0 cf. above)

$$\text{(1)} \qquad C\omega(2\eta_0) \leqq \gamma_0\varepsilon.$$

Then choose a P (with $1 \in P$) so that all

$$\text{(2)} \qquad \mathrm{Osc}_P(f)\,\mathrm{Osc}_P(f_0), \mathrm{Osc}_P(f_1), \ldots, \mathrm{Osc}_P(f_t) \leqq \eta_0.$$

(Use Lemma I in §8.) Finally form as in (2) in §18 (remember also Lemma IV in §8)

$$\text{(3)} \qquad O^{\circledast} = O^{\circledast}(C, P) \leqq P \leqq O_*.$$

Now consider two O', O'' with

$$\text{(4)} \qquad 1 \in O', O'' \leqq O_0 \cdot O^{\circledast}$$

(for O_0 cf. above). Then Definitions I, II in §16 give

$$\text{(5)} \qquad \left.\begin{aligned} \mathcal{M}_{O'} &= \mathcal{M}'_{O'} \\ \mathcal{M}_{O''} &= \mathcal{M}'_{O''} \end{aligned}\right\} \quad \text{following Definition I in §16.}$$

Consider further two means

$$\text{(6)} \qquad \mathbb{M}'(h) \quad \text{in } \mathcal{M}_{O'}, \qquad \mathbb{M}''(h) \quad \text{in } \mathcal{M}_{O''}.$$

Clearly Lemma III in §15 (compare it with Definition I in §16) applies to both of them. Therefore they fulfill (3)–(5) in §18: (3) in §18 owing to (β) in Lemma III in §15, (4) in §18 owing to (γ) eod., and (5) in §18 owing to (α) eod. together with the lemma in §17—or its corollary.

Therefore Lemma II in §19 applies to f_1, f_0 (in place of its f, g_1). Our (2) gives

$$\text{(7)} \qquad \eta \leqq \eta_0$$

for the η loc. cit. Now (β) loc. cit. gives

$$|\mathbb{M}'(f)\mathbb{M}''(f_0) - \mathbb{M}''(f)\mathbb{M}'(f_0)| \leqq \omega(2\eta); \tag{8}$$

hence by (7) and (1)

$$|\mathbb{M}'(f)\mathbb{M}'''(f_0) - \mathbb{M}''(f)\mathbb{M}'(f_0)| \leqq \gamma_0\varepsilon. \tag{9}$$

And (δ) in Lemma III in §15 gives

$$\mathbb{M}'(f_0) = \mathbb{M}''(f_0) = \gamma_0$$

whereby (9) becomes

$$|\mathbb{M}'(f) - \mathbb{M}''(f)| \leqq \varepsilon. \tag{10}$$

Summing up (4) (for O', O'') and (6) imply (10). This means, however, that the condition (β) of Lemma I in §14 is satisfied, if we choose its $O_1 = O_1(f, \varepsilon) = O_0 \cdot O^{\circledast}$ (cf. our (3) and (4)). This is legitimate since our O_0 depends on C only (cf. Definition I in §16 and Lemma II in §15), and since our $O^{\circledast}$ depends on C, P only (cf. the beginning of §18 in particular (2) eod.), while P in turn depends on C, f, ε only (cf. our (1), (2)).

Thus the proof is completed.

The significant results of our discussions can now be enunciated in one theorem:

Theorem. *Given a C, with $C^i \neq \theta$, these are true:*

(α) *There exists one and only one mean $\overline{\mathbb{M}}(f)$ which fulfills* (γ) *in Lemma* I *in* §14 *for the sets $\mathcal{M}_O$ of Definitions* I, II *in* §16.

(β) *This $\overline{\mathbb{M}}(f)$ is l.i. and not identically* 0.

(γ) *There exists one and only one l.i. measure $\tau(M)$ in G, which is the correspondent of the above $\overline{\mathbb{M}}(f)$ in the sense of the definition in* §13, *i.e. with*

$$\overline{\mathbb{M}}(f) = \int_G f(x)\, d\tau_x \quad \textit{for all } f \in \mathcal{F}_C.$$

This $\tau(M)$ is not identically 0.

PROOF. Ad (α): Immediate by Lemma I above and Lemma I in §14.

Ad (β): Considering Lemma I above, $\overline{\mathbb{M}}(f)$ is l.i. by Lemma I in §14, and it is not identically 0 by Lemma II in §14.

Ad (γ): Immediate by our (β) and by the definition and Lemma II in §13.

Thus the existence of Haar's measure is established by a purely constructive process.

The remarks made at the end of §16 may be recalled once more. The equidistributed sets, which make up the $\mathcal{M}_O$ (cf. α) in our above theorem (i.e. cf. Definitions I, II in §16) can be obtained with the help of the lemmas of §5 and of §6—or equivalently by the discussion of §7.

Section 5 deals with "minimum coverings". Section 6 deals with "maximum fillings", i.e. with "closest packings". These notions lead therefore to Haar's measure—and this is every (locally compact) group G.