Lecture Notes in Mathematics

A collection of informal reports and seminars
Edited by A. Dold, Heidelberg and B. Eckmann, Zürich

133

Flemming Topsøe

Department of Mathematics, University of Copenhagen

Topology and Measure

Springer-Verlag
Berlin · Heidelberg · New York 1970

 Library of Congress Catalog Card Number 75-120379. Printed in Germany.
Title No. 3289.

CONTENTS

Preface introduction acknowledgments IV
Preliminaries IX

PART I

1. Measure and integral, definitions 1
2. Basic result on construction of a measure 3
3. Basic result on construction of an integral 6
4. Finitely additive theory 15
5. From "Baire" measures to "Borel" measures, an abstract approach 21
6. Construction of measures by approximation from outside and by approximation from inside 26
7. On the possibility of providing a space of measures with a vague topology 31

PART II

8. Definition and basic properties of the weak topology 40
9. Compactness in the weak topology 42
10. Criteria for weak convergence 45
11. On the structure of $\mathcal{M}_+(X)$ 47
12. A problem related to questions of uniformity 51
13. First solution of the ξ-problem 54
14. Second solution of the ξ-problem 60
15. Uniformity classes 64
16. Joint continuity 66
17. Preservation of weak convergence 68

Notes and remarks 72

References 78

PREFACE INTRODUCTION ACKNOWLEDGMENTS

Below we shall comment on the development which led to the results of the present volume (and of [26]).

It will be seen that our investigations took their starting point in the theory of weak convergence of (probability-) measures, and that the results on measure and integration theory to be found in part I emerged as a kind of "by-product".

During the inspiring lectures of professor P. Billingsley, I was for the first time presented to the theory of weak convergence. These lectures, which were based on the manuscript to the book [3], took place in 1964-65 while professor Billingsley visited the institute of math. statistics at the university of Copenhagen. The contact with Billingsley resulted, among other things, in the joint paper [4]. This paper was the starting point of the development leading to the results obtained in sections 12-17.

The academic year 1965-66 I spent at the statistical laboratory, the university of Cambridge, England. There I met professor K.R. Parthasarathy, who was at that time working on his book [17], dealing with weak convergence too. Once, professor Parthasarathy posed the following question and explained the significance of the problem: Let X be a Polish space, $\mathcal{A}$ a field of subsets of X generating the Borel σ-field and P, $(P_n)_{n\geq 1}$ probability measures on X such that lim $P_n A = PA$ for all A in $\mathcal{A}$; is it then true, or under what additional assumptions is it true, that P_n converges weakly to P?

A solution to the problem of Parthasarathy was published in [24]. It was quite obvious that the reasoning had very little of "Polishness" in it and, with this motivation, we became interested in developing a theory of weak convergence of tight measures on arbitrary topological Hausdorff spaces. Of course, this had already been done by Varadarajan in 1961 and by his predecessors; however, despite the great achievements in this previous research, there were some points which seemed unsatisfactory to me (notably, the choice of σ-fields).

One of the first results we were able to obtain was that $\mathcal{M}_+(X;t)$ in its weak topology is a Hausdorff space for any Hausdorff space X (Theorem 11.2, (i)). This innocent looking result was in fact enough to convince us that we were working with a fruitful definition of weak convergence. Note, that the definition is based on semi-continuous functions and not on continuous functions. The definition and further study of the weak topology, was also influenced by the announcement [20] of L.Schwartz of a theory of Radon measures (= tight measures) on arbitrary Hausdorff spaces. The following remark cited from [20] was helpful: "All the properties of Radon measures given in Bourbaki extend here, provided continuous functions don't occur in their statement (but semi-continuous functions may occur)".

Clearly, the theory of weak convergence would be worthless, could we not establish suitable criteria for compactness. Led by certain observations (which are at present uninteresting) we came to the conviction, that Prohorovs condition of tightness would also be sufficient in the general case. I had a very specific idea as to how the hoped-for result could be proved but couldn't push it through. At that time (1968) I received affirmative solutions of the problem from P.A. Meyer and from L. Schwartz. Both used the idea - which never occured to me - of reducing the problem to the compact case.

The problem was settled, but for some reasons, perhaps stubbornness too, I was not fully satisfied and tried again - in vain - to push my own ideas through. Then a paper by J. Kisyński ([14]) was published which did just what we wanted to. Indeed, it turned out that with Kisyński's result at hand we could obtain necessary and sufficient conditions for compactness in $\mathcal{M}_+(X;t)$. However, the condition we arrived at (cf. Theorem 9.1) was rather complicated and we tried to derive simpler ones. We did nothave much luck with the tightness condition (see notes and remarks to section 9), and we started to look into the condition of τ-smoothness. Then one should work in the space $\mathcal{M}_+(X;\tau)$ rather than in the space $\mathcal{M}_+(X;t)$, and the compact sets in X did no longer play the dominant role, but was to some extent taken over by the closed sets. It was natural to examine once more the proof of Kisyński's theorem and see if one could axiomatize it so that it would cover the two cases. This was in fact quite easy to do (Theorem 2.2), and as a consequence we obtained the desired result on τ-smoothness and compactness (Theorem 9.2)

It turned out that Theorem 2.2 also contained the extension theorem in abstract measure theory known as Carathéodory's theorem (see [26]). Thus we have an instance of a result that can be used as the key to topological measure theory as well as to abstract measure theory. We found this point worth while exploring in its own right and, for that reason, we established an analogous result, Theorem 3.13, which is an extension theorem for integrals. This result contains a version of Daniells theorem. We do hope that these abstract results will turn out to be of some interest, also in the teaching of measure and integration theory. We are fully aware that we have contributed only a modest amount to the ideas in the proofs of Theorems 2.2 and 3.13. The main idea is still the brillant idea from 1914 due to C. Carathéodory (see [7]).

We have tried to round off the results on measure theory by inclusion of sections 4,5 and 6. Section 4 deals with finitely additive measures and integrals, and contains a version, due to A. Markoff, of the "Riesz representation theorem".

Provoked by a stimulating discussion with E.T. Kehlet, we included a section explaining, in the framework of the present theory, why it is possible, in locally compact spaces, to construct regular Borel-measures from certain set-functions only defined on the compact Baire-sets.

The main results in section 6 on construction of measures by approximation from outside or from inside are in fact vital for the solution of the original problem on compactness in the topology of weak convergence. Compactness problems in spaces of measures provided with other topologies than that of weak convergence can also be handled by appealing to the results of section 6 - one instance of this is demonstrated in the following section dealing with the vague topology, and yet another topology, much stronger than the ones already mentioned, is discussed in our paper [26].

It is probably true, that all small tricks employed in the first six sections can be found scattered in the huge literature of measure and integration theory. I have made no attempt to make reference to this literature - for one reason, I do not know it well enough, and, for another, it seems as if the main idea, viz. the idea to base the theory of measure on inner measure, has not been taken up previously. There is, however, at least one exception: As pointed out to us by S.D. Chatterji, the paper [21] from 1955 by Srinivasan works explicitely with inner measure. Also, it may be true that some of the results to be found in part I (notably Theorem 2.2) are contained in the paper [18] from 1951 by Pettis.

We have now described how the work was carried out and what the motivations were. In the text we have of course arranged the material in a more systematic order. There are two parts. Part II deals with weak convergence andpart I contains the material not directly connected with weak convergence. The main theme of part I is construction of measures and integrals. A section of notes and remarks is included at the end of the text.

The reader, who is not interested in going through the entire text may find the following suggestions for reading attractive: Sections 1,2, 3 (omitting all proofs), 5(again omitting proofs), 6 (only paying attention to Theorem 6.2), 8,9,10, and then, at last, the reader has to decide how much of sections 12-17 he wants to go through; this is perhaps best done by looking into sections 15-17 where the results of sections 12-14 are applied, but even if the applications are found to be worth while, the reader may find it unbearable to go through the proofs of sections 13 and 14. However, in order to understand what is going on, it is quite sufficient to acquire familiarity with the proofs of the much simpler Theorem 2 of Billingsley and Topsøe [4] and of Theorem 2 of Topsøe [22]. It can be said that once the problem in the more complicated set-up has been properly formulated (cf. section 12), it is a matter of routine to solve it, knowing the ideas of the above mentioned papers.

It follows from what is said above, that the persons I am mostly indebted to are professor Billingsley and professor Parthasarathy, and I want to express my sincere thanks.

PRELIMINARIES

P1. $\hat{R}$ denotes the reals and $\hat{R}_+$ the non-negative reals.

P2. Sometimes we find it convinient to call a non-empty class of subsets of a set X, a paving on X. A $(\cup f, \cap c)$-paving is a paving closed under finite unions and countable intersections. If in addition the empty set $\emptyset$ is a member of the paving, we speak of a $(\emptyset, \cup f, \cap c)$-paving. This method of notation is employed systematically through the text. The paving of all subsets of X is denoted $\mathfrak{D}(X)$.

P3. A class $\mathcal{A} \subseteq \mathfrak{D}(X)$ is filtering to the left or filtering downwards if, to any pair (A_1, A_2) of sets in $\mathcal{A}$ we can find $A \in \mathcal{A}$ such that $A \subseteq A_1 \cap A_2$. We write $\mathcal{A} \downarrow A_0$ if $\mathcal{A}$ is filtering to the left and $A_0 = \bigcap\{A \mid A \in \mathcal{A}\}$. A class $\mathcal{F}$ of functions $f: X \to \hat{R}$ filters to the left if $f_1, f_2 \in \mathcal{F} \Rightarrow \exists f \in \mathcal{F} : f \leq \min(f_1, f_2)$. Filtering to the right (or upwards) is defined analogously.

P4. We shall work with nets rather than with filters. Our prefered labels for directed sets are D, E and I. Elements of a directed set denoted by the letter D[I] will always be denoted by the letter α[i]. Examples: $(x_\alpha)_{\alpha \in D}$, $(f_i)_{i \in I}$ or just (x_α), (f_i). We write $x_\alpha \in A$, eventually or just $x_\alpha \in A$, ev. if for some $\alpha_0 \in D$ we have $x_\alpha \in A$ for all $\alpha \geq \alpha_0$. We write $x_\alpha \in A$, frequently or just $x_\alpha \in A$, freq. if for every $\alpha \in D$ we have $x_\beta \in A$ for some $\beta \geq \alpha$.

P5. A notion of convergence $\xrightarrow{t}$ on a set X is a class of pairs (ξ, x) where ξ is a net on X and x a point of X such that certain conditions are fulfilled. We write $x_\alpha \xrightarrow{t} x$ to indicate that the pair $((x_\alpha)_{\alpha \in D}, x)$ is a member of the class. We require that the following conditions are fulfilled:

(i): $x_\alpha = x$ for all $\alpha \in D \Rightarrow x_\alpha \to_\tau x$,

(ii): $x_\alpha \to_\tau x \Rightarrow x_{\alpha_\beta} \to_\tau x$ for every subnet (x_{α_β}) of (x_α),

(iii): $x_\alpha \to_\tau x,\ x_\alpha \to_\tau y \Rightarrow x = y$,

(iv): If $(x_\alpha)_{\alpha\in D}$, x is such that every subnet of (x_α) contains a further subnet converging to x, then the net (x_α) itself converges to x.

P6. We shall assume that all our basic topological spaces are Hausdorff spaces. Thus every topological space denoted by the letter X is assumed to be a Hausdorff space. For a topological space X we denote by $\mathcal{F}(X)$, $\mathcal{K}(X)$, $\mathcal{G}(X)$ and $\mathcal{B}(X)$ the pavings on X of closed, compact, open and Borel sets, respectively. Sets denoted by the letters F,K,G are usually assumed without further mentioning to be closed, compact or open, respectively.

P7. A net (x_α) on a topological space X is said to be <u>compact</u> if every subnet has a further subnet which converges (or, equivalently, if every universal subnet of (x_α) converges). A subset A of X is called <u>net-compact</u> if every net on A has a convergent subnet (or, equivalently, if every universal net on A converges). In case X is a regular topological space, $A \subseteq X$ is net-compact if and only if A is relatively compact.

P8. Let X be a topological space and (F_α) a net on $\mathcal{F}(X)$. We define two sets F_* and F^*, both closed sets, by

$$F_* = \{x \mid \forall_{N(x)}\ N(x)\cap F_\alpha \neq \emptyset,\ \text{ev.}\},$$

$$F^* = \{x \mid \forall_{N(x)}\ N(x)\cap F_\alpha \neq \emptyset,\ \text{freq.}\}.$$

Here, N(x) denotes some neighbourhood of x. If $F_* = F^*$, we write $F_\alpha \to F_*$ and we say that F_α <u>converges in the notion of closed topological convergence</u> to F_*. This notion is indeed a notion of convergence on $\mathcal{F}(X)$, and it has the interesting property (Hausdorffs selection theorem) that every net on $\mathcal{F}(X)$ has a convergent subnet. The notion is topological if and only if X is locally compact.

P9. For X a topological space and f a function $X \to \hat{\mathbb{R}}$ (say bounded) we define the lower semi-continuous envelope f_* and the upper semi-continuous envelope f^* of f by

$$f_* = \sup\{g \mid g \leq f,\ g \text{ l.s.c.}\},$$

$$f^* = \inf\{g \mid g \geq f,\ g \text{ u.s.c.}\}.$$

f_* is l.s.c., f^* is u.s.c., $f_* \leq f \leq f^*$ holds, and, furthermore, we have

$$D(f) = \{f_* < f^*\},$$

where D(f) denotes the set of discontinuity points of f. It follows that

$$D(f) = \bigcup_1^\infty \{f^* - f_* \geq 1/n\}$$

is a F_σ-set, in particular a Borel-set.

P10. Let X be a topological space. A class $\mathcal{A} \subseteq \hat{D}(X)$ is said to <u>separate points</u> T_2 if to any pair (x,y) of distinct points in X there exists $A \in \mathcal{A}$ such that $x \in \mathring{A}$ and $y \notin \bar{A}$. $\mathcal{A}$ is said to <u>separate points and closed sets</u> T_1 if $x \notin F$, $F \in \mathcal{F}(X) \Rightarrow x \in \mathring{A}$, $A \cap F = \emptyset$ for some $A \in \mathcal{A}$. $\mathcal{A}$ is said to <u>separate points and closed sets</u> T_2 if $x \notin F$, $F \in \mathcal{F}(X)$ $\Rightarrow x \in \mathring{A}$, $\bar{A} \cap F = \emptyset$ for some $A \in \mathcal{A}$. If $\mathcal{A}$ separate points and closed sets T_1 and if X is regular then $\mathcal{A}$ separate points and closed sets T_2.

P11. Let X be an abstract set and $\mathcal{A}$ and $\mathcal{S}$ two classes of subsets of X such that: $\mathcal{A} \subseteq \mathcal{S}$, $\mathcal{A}$ is closed under finite intersections, the complement of every set in $\mathcal{S}$ is in $\mathcal{S}$, the union of two disjoint sets in $\mathcal{S}$ is in $\mathcal{S}$, and $S_1 \setminus S_2 \in \mathcal{S}$ whenever $S_1, S_2 \in \mathcal{S}$ and $S_1 \supseteq S_2$. Then $\alpha(\mathcal{A})$, the algebra spanned by $\mathcal{A}$, is contained in $\mathcal{S}$.

P12. By a <u>set-function</u> we mean a non-negative, possibly infinite valued function defined on a paving. Let β be a set-function defined on the paving $\mathcal{A}$. In the definitions below we only require the defining relations to hold when they make sense. β is <u>monotone</u> if $A_1 \subseteq A_2 \Rightarrow \beta A_1 \leq \beta A_2$.

β is <u>subadditive</u> if $\beta(A_1 \cup A_2) \leq \beta A_1 + \beta A_2$ holds. β is <u>additive</u> if $A_1 \cap A_2 = \emptyset \Rightarrow \beta(A_1 \cup A_2) = \beta A_1 + \beta A_2$. β is <u>modular</u> if $\emptyset \in \mathcal{A}$, if $\beta\emptyset = 0$, and if $\beta(A_1 \cup A_2) + \beta(A_1 \cap A_2) = \beta A_1 + \beta A_2$ holds. A monotone set function β defined on $\mathcal{A}$ is σ-<u>smooth</u> [τ-<u>smooth</u>] <u>with respect to the paving</u> $\mathcal{K}$ if, for any countable [arbitrary] subclass $\mathcal{K}^*$ of $\mathcal{K}$ which filters downwards to a set A_0 in $\mathcal{A}$ ($\mathcal{K}^* \downarrow A_0$), we have

$$\beta A_0 = \inf\{\beta A \mid A \supseteq K^* \text{ for some } K^* \in \mathcal{K}^*\},$$

provided the right hand side in this equation is finite. If $\emptyset \in \mathcal{A}$ and if we only require the last relation to hold when $A_0 = \emptyset$, then we obtain the definition of set-functions which are σ-<u>smooth at</u> $\emptyset$ [τ-<u>smooth at</u> $\emptyset$] <u>w.r.t.</u> $\mathcal{K}$. If $\mathcal{K} = \mathcal{A}$ in the last definitions, we call the set-function σ-<u>smooth</u>, τ-<u>smooth</u>, σ-<u>smooth at</u> $\emptyset$, or τ-<u>smooth at</u> $\emptyset$, respectively. β is <u>regular w.r.t. the paving</u> $\mathcal{K}$ if $\mathcal{K} \subseteq \mathcal{A}$ and if

$$\beta A = \sup\{\beta K \mid K \subseteq A,\ K \in \mathcal{K}\}$$

holds. β is <u>tight</u> if β is finite and if, whenever $A_1 \supseteq A_2$, the relation

$$\sup\{\beta A \mid A \subseteq A_1 \setminus A_2\} = \beta A_1 - \beta A_2$$

holds. β is a <u>content</u> if $\mathcal{A}$ is a $(\emptyset, \cup f, \cap f)$ paving and if β is finite, monotone, subadditive and additive.

P13. Let X be a topological space. By $\mathcal{M}_+(X)$ we denote the space of all non-negative totally finite measures defined on $\mathcal{B}(X)$. $\mu \in \mathcal{M}_+(X)$ is <u>regular</u> if μ is regular w.r.t. the paving $\mathcal{F}(X)$, and μ is <u>tight</u>, or a <u>Radon measure</u>, if μ is regular w.r.t. the paving $\mathcal{K}(X)$. $\mu \in \mathcal{M}_+(X)$ is τ-<u>smooth</u> if

$$\mu\Big(\bigcap_{F \in \mathcal{F}} F\Big) = \inf_{F \in \mathcal{F}} \mu F$$

holds for any family $\mathcal{F}$ of closed sets filtering to the left. $\mathcal{M}_+(X;r)$, $\mathcal{M}_+(X;\tau)$, $\mathcal{M}_+(X;t)$, $\mathcal{M}_+(X;r,\tau)$ denote the sets of regular, τ-smooth, tight, and regular τ-smooth measures in $\mathcal{M}_+(X)$, respectively.

P14. If $A \in \mathcal{B}(X)$ and $\mu \in \mathcal{M}_+(X;t)$ then $\mu|A$, the restriction of μ to A, is a measure in $\mathcal{M}_+(X,t)$ too. An analogous statement holds for $\mathcal{M}_+(X;r,\tau)$.

P15. If $\mu \in \mathcal{M}_+(X;t)$ then $\mu \in \mathcal{M}_+(X;r,\tau)$. If $\mu \in \mathcal{M}_+(X;\tau)$ and if $\mathcal{F}$ is a uniformly bounded family of u.s.c. functions $X \to \hat{R}$ filtering to the left then we have

$$\mu(\inf_{f\in\mathcal{F}} f) = \inf_{f\in\mathcal{F}} \mu(f).$$

This is easily proved by applying the inequality

$$\frac{1}{k}\sum_{1}^{k}\mu(\{g \geq \tfrac{\nu}{k}\}) \leq \int g\,d\mu \leq \frac{\mu X}{k} + \frac{1}{k}\sum_{1}^{k}\mu(\{g \geq \tfrac{\nu}{k}\})$$

valid for any $k \geq 1$ and for any measurable function $g: X \to \hat{R}$ with $0 < g < 1$.

P16. If every closed subset of X is a G_δ-set than every measure (in $\mathcal{M}_+(X)$) is regular.

If X is fully Lindelöf (i.e. every family of open sets has a countable subfamily with the same union), then every measure is τ-smooth.

If X can be metrized with a complete metric, or if X is locally compact, then every τ-smooth measure is tight.

If X is regular and μ τ-smooth then we have

$$\mu G_0 = \sup\{\mu G \mid \bar{G} \subseteq G_0\};\ G_0 \in \mathcal{G}(X),$$

$$\mu F_0 = \inf\{\mu F \mid \mathring{F} \supseteq F_0\};\ F_0 \in \mathcal{F}(X).$$

In particular, it follows that μ is regular.

If X is completely regular and μ τ-smooth then

$$\mu G_0 = \sup\{\mu G \mid G \subseteq G_0,\ \mu(\partial G) = 0\}.$$

$$= \sup\{\mu F \mid F \subseteq G_0,\ \mu(\partial F) = 0\};\quad G_0 \in \mathcal{G}(X).$$

$$\mu F_0 = \inf\{\mu F \mid F \supseteq F_0,\ \mu(\partial F) = 0\}.$$

$$= \inf\{\mu G \mid G \supseteq F_0,\ \mu(\partial G) = 0\};\ F_0 \in \mathcal{F}(X).$$

P17. A subset F of X is called <u>support</u> for $\mu \in \mathcal{M}_+(X)$, and we write $F = \mathrm{supp}(\mu)$, if $F \in \mathcal{F}(X)$, if $\mu F = \mu X$ and if $F' \in \mathcal{F}(X)$, $\mu F' = \mu X$ $\Rightarrow F' \supseteq F$. $\mathrm{supp}(\mu)$ is uniquely determined.

Every τ-smooth measure has a support.

P18. A function $f: X \to \hat{R}$ is called a μ-continuity function if f is measurable (w.r.t. $\mathcal{B}(X)$) and if $\mu(D(f)) = 0$. The set $A \in \mathcal{B}(X)$ is a μ-continuity set if $\mu(\partial A) = 0$.

P19. If $\mathcal{A}$ is a $(\cup f, \cap f)$-paving on X separating points T_2 and if μ_1 and μ_2 are two tight measures such that

$$\mu_1(\mathring{A}) \leq \mu_2(\bar{A}); \quad A \in \mathcal{A}$$

then $\mu_1 \leq \mu_2$. If, in addition $\mu_1 X = \mu_2 X$, then $\mu_1 = \mu_2$.

If X is regular and $\mathcal{A}$ a $(\cup f)$-paving on X separating points and closed sets T_2 and if μ_1 and μ_2 are measures in $\mathcal{M}_+(X;\tau)$ such that

$$\mu_1(\mathring{A}) \leq \mu_2(\bar{A}); \quad A \in \mathcal{A}$$

then $\mu_1 \leq \mu_2$. If, in addition, $\mu_1 X = \mu_2 X$, then $\mu_1 = \mu_2$.