Graduate Texts in Mathematics

John L. Kelley T. P. Srinivasan

Measure and Integral

Volume 1

测度和积分 第1卷 [英]

Springer-Verlag
World Publishing Corp

John L. Kelley   T. P. Srinivasan

Measure and Integral

Volume 1

Springer-Verlag
World Publishing Corp

John L. Kelley
Department of Mathematics
University of California
Berkeley, CA 94720
U.S.A.

T.P. Srinivasan
Department of Mathematics
The University of Kansas
Lawrence, KN 66045
U.S.A.

AMS Classification: 28-01

Library of Congress Cataloging-in-Publication Data
Kelley, John L.
Measure and integral/John L. Kelley, T.P. Srinivasan.
p. cm.—(Graduate texts in mathematics; 116)
Bibliography: p.
Includes index.
ISBN 0-387-96633-1
1. Measure theory. 2. Integrals, Generalized. I. Srinivasan, T.P.
II. Title. III. Series.
QA312.K44 1988
515.4'2—dc19 87-26571

Reprinted by World Publishing Corporation, Beijing, 1990
for distribution and sale in The People's Republic of China only

ISBN 7-5062-0739-7

ISBN 0-387-96633-1 Springer-Verlag New York Berlin Heidelberg
ISBN 3-540-96633-1 Springer-Verlag Berlin Heidelberg New York

PREFACE

This is a systematic exposition of the basic part of the theory of measure and integration. The book is intended to be a usable text for students with no previous knowledge of measure theory or Lebesgue integration, but it is also intended to include the results most commonly used in functional analysis. Our two intentions are some what conflicting, and we have attempted a resolution as follows.

The main body of the text requires only a first course in analysis as background. It is a study of abstract measures and integrals, and comprises a reasonably complete account of Borel measures and integration for $\mathbb{R}$. Each chapter is generally followed by one or more supplements. These, comprising over a third of the book, require somewhat more mathematical background and maturity than the body of the text (in particular, some knowledge of general topology is assumed) and the presentation is a little more brisk and informal. The material presented includes the theory of Borel measures and integration for $\mathbb{R}^n$, the general theory of integration for locally compact Hausdorff spaces, and the first dozen results about invariant measures for groups.

Most of the results expounded here are conventional in general character, if not in detail, but the methods are less so. The following brief overview may clarify this assertion.

The first chapter prepares for the study of Borel measures for $\mathbb{R}$. This class of measures is important and interesting in its own right and it furnishes nice illustrations for the general theory as it develops. We begin with a brief analysis of length functions, which are functions on the class $\mathscr{I}$ of closed intervals that satisfy three axioms which are eventually shown to ensure that they extend to measures. It is shown in chapter 1 that every length function has a unique extension λ to the lattice $\mathscr{L}$ of sets generated by $\mathscr{J}$ so that λ is *exact*, in the sense that $\lambda(A) = \lambda(B) + sup\{\lambda(C): C \in \mathscr{L}$ and $C \subset A \setminus B\}$ for members A and B of $\mathscr{L}$ with $A \subset B$.

The second chapter details the construction of a pre-integral from a pre-measure. A real valued function μ on a family $\mathscr{A}$ of sets that is closed under finite intersection is a *pre-measure* iff it has a countably additive non-negative extension to the ring of sets generated by $\mathscr{A}$ (e.g., an exact function μ that is continuous at $\varnothing$). Each length function is a pre-measure. If μ is an exact function on $\mathscr{A}$, the map $\chi_A \mapsto \mu(A)$ for A in $\mathscr{A}$ has a linear extension I to the vector space L spanned by the characteristic functions χ_A, and the space L is a vector lattice with truncation: $I \wedge f \in L$ if $f \in L$. If μ is a pre-measure, then the positive linear functional I has the property: if $\{f_n\}_n$ is a decreasing sequence in L that converges pointwise to zero, then $lim_n I(f_n) = 0$. Such a functional I is a *pre-integral*. An *integral* is a pre-integral with the Beppo Levi property: if $\{f_n\}_n$ is an increasing sequence in L converging pointwise to a function f and $sup_n I(f_n) < \infty$, then $f \in L$ and $lim_n I(f_n) = I(f)$.

In chapter 3 we construct the Daniell–Stone extension L^1 of a pre-integral I on L by a simple process which makes clear that the extension is a completion under the L^1 norm $\|f\|_1 = I(|f|)$. Briefly: a set E is called *null* iff there is a sequence $\{f_n\}_n$ in L with $\sum_n \|f_n\|_1 < \infty$ such that $\sum_n |f_n(x)| = \infty$ for all x in E, and a function g belongs to L^1 iff g is the pointwise limit, except for the points in some null set, of a sequence $\{g_n\}_n$ in L such that $\sum_n \|g_{n+1} - g_n\|_1 < \infty$ (such sequences are called *swiftly convergent*). Then L^1 is a norm completion of L and the natural extension of I to L^1 is an integral. The methods of the chapter, also imply for an arbitrary integral, that the domain is norm complete and the monotone convergence and the dominated convergence theorems hold. These results require no measure theory; they bring out vividly the fundamental character of M. H. Stone's axioms for an integral.

A *measure* is a real (finite) valued non-negative countably additive function on a δ-ring (a ring closed under countable intersection). If J is an arbitrary integral on M, then the family $\mathscr{A} = \{A: \chi_A \in M\}$ is a δ-ring and the function $A \mapsto J(\chi_A)$ is a measure, the measure induced by the integral J. Chapter 4 details this procedure and applies the result, together with the pre-measure to pre-integral to integral theorems of the preceding chapters to show that each exact function that is continuous at $\varnothing$ has an extension that is a measure. A supplement presents the standard construction of regular Borel measures and another supplement derives the existence of Haar measure.

A measure μ on a δ-ring $\mathscr{A}$ is also a pre-measure; it induces a pre-integral, and this in turn induces an integral. But there is a more direct way to obtain an integral from the measure μ: A real valued function f belongs to $L_1(\mu)$ iff there is $\{a_n\}_n$ in $\mathbf{R}$ and $\{A_n\}_n$ in $\mathscr{A}$ such that

$\sum_n |a_n| \mu(A_n) < \infty$ and $f(x) = \sum_n a_n \chi_{A_n}(x)$ for all x, and in this case the *integral* $I_\mu(f)$ is defined to be $\sum_n a_n \mu(A_n)$. This construction is given in chapter 6, and it is shown that every integral is the integral with respect to the measure it induces.

Chapter 6 requires facts about measurability that are purely set theoretic in character and these are developed in chapter 5. The critical results are: Call a function f $\mathscr{A}$ *σ-simple* (or $\mathscr{A}$ *σ^+-simple*) iff $f = \sum_n a_n \chi_{A_n}$ for some $\{A_n\}_n$ in $\mathscr{A}$ and $\{a_n\}_n$ in $\mathbb{R}$ (in $\mathbb{R}^+$, respectively). Then, if $\mathscr{A}$ is a δ-ring, a real valued function f is $\mathscr{A}$ σ-simple iff it has a support in $\mathscr{A}_\sigma$ and is locally $\mathscr{A}$ measurable (if B is an arbitrary Borel subset of $\mathbb{R}$, then $A \cap f^{-1}[B]$ belongs to $\mathscr{A}$ for each A in $\mathscr{A}$). Moreover, if such a function is non-negative, it is $\mathscr{A}$ σ^+-simple.

Chapter 7 is devoted to product measures and product integrals. It is concerned with conditions that relate the integral of a function f w.r.t. $\mu \otimes \nu$ to the iterated integrals $\int (\int f(x,y)\, d\mu x)\, d\nu y$ and $\int (\int f(x,y)\, d\nu y)\, d\mu x$. We follow the natural approach, deriving the Fubini theorem from the Tonelli theorem, and the latter leads us to grudgingly allow that some perfectly respectable σ-simple functions have infinite integrals (we call these functions *integrable in the extended sense*, or *integrable**).

Countably additive non-negative functions μ to the extended set $\mathbb{R}^*$ of reals (*measures in the extended sense* or *measures**) also arise naturally (chapter 8) as images of measures under reasonable mappings. If μ is a measure on a σ-field $\mathscr{A}$ of subsets of X, $\mathscr{B}$ is a σ-field for Y, and $T: X \to Y$ is $\mathscr{A} - \mathscr{B}$ measurable, then the image measure $T\mu$ is defined by $T\mu(B) = \mu(T^{-1}[B])$ for each B in $\mathscr{B}$. If $\mathscr{A}$ is a δ-ring but not a σ-field, there is a possibly infinite valued measure that can appropriately be called the T image of μ. We compute the image of Borel–Lebesgue measure for $\mathbb{R}$ under a smooth map, and so encounter indefinite integrals.

Indefinite integrals w.r.t. a σ-finite measure μ are characterized in chapter 9, and the principal result, the Radon–Nikodym theorem, is extended to decomposable measures and regular Borel measures in a supplement. Chapter 10 begins the study of Banach spaces. The duals of some standard spaces are characterized, and in a supplement our methods are used to establish very simply, or at least σ-simply, the basic facts about Bochner integrals.

This book is based on various lectures given by one or the other of us in 1965 and later, at the Indian Institute of Technology, Kanpur; Panjab University, Chandigarh; University of California, Berkeley; and the University of Kansas. We were originally motivated by curiosity about how a σ-simple approach would work; it did work, and a version of most of this text appeared as preprints in 1968, 1972 and 1979, under the title "Measures and Integrals." Since that time our point of view has changed on several matters (but not on σ-simplicity) and the techniques have been refined.

This is the first of two volumes on *Measure and Integral*. The exercises, problems, and additional supplements will appear as a companion volume to be published as soon as we can sift and edit a large disorganized mass of manuscript.

We are grateful to Klaus Bichteler, Harlan Glaz, T. Parthasarathy, and Allan Shields for suggestions and criticisms of earlier versions of this work and to Dorothy Maharam Stone and I. Namioka for their review of the final manuscript. We are indebted to our students for their comments and their insights. We owe thanks to Jean Steffey, Judy LaFollette, Carol Johnson, and especially to Ying Kelley and Sharon Gumm for assistance in preparation of the manuscript, and to Saroja Srinivasan for her nonmeasurable support.

This work was made possible by support granted at one time or another by the Miller Foundation of the University of California, Berkeley, the National Science Foundation, the Panjab (India) University, and the University of Kansas. We thank them.

J. L. KELLEY T. P. SRINIVASAN

CONTENTS

Chapter 0

PRELIMINARIES

This brief review of a few conventions, definitions and elementary propositions is for reference to be used as the need arises.

SETS

We shall be concerned with sets and with the membership relation, $\in$. If A and B are sets then $A = B$ iff A and B have the same members; i.e., for all x, $x \in A$ iff $x \in B$. A set A is a subset of a set B (B is a superset of A, $A \subset B$, $B \supset A$) iff $x \in B$ whenever $x \in A$. Thus $A = B$ iff $A \subset B$ and $B \subset A$. The empty set is denoted $\varnothing$.

If A and B are sets then the union of A and B is $A \cup B$, $\{x: x \in A \textit{ or } x \in B\}$; the intersection $A \cap B$ is $\{x: x \in A \textit{ and } x \in B\}$; the difference $A \setminus B$ is $\{x: x \in A \textit{ and } x \notin B\}$; the **symmetric difference** $A \triangle B$ is $(A \cup B) \setminus (A \cap B)$; and the Cartesian product $A \times B$ is $\{(x, y): x \in A, y \in B\}$. The operations of union, intersection, and symmetric difference are commutative and associative, $\cap$ distributes over $\cup$ and $\triangle$, and $\cup$ distributes over $\cap$. The set $\varnothing$ is an identity for both $\cup$ and $\triangle$.

If, for each member t of an index set T, A_t is a set, then this correspondence is called an **indexed family**, or sometimes just a **family** of sets and denoted $\{A_t\}_{t \in T}$. The union of the members of the family is $\bigcup_{t \in T} A_t = \bigcup \{A_t: t \in T\} = \{x: x \in A_t \textit{ for some member } t \textit{ of } T\}$ and the intersection is $\bigcap_{t \in T} A_t = \bigcap \{A_t: t \in T\} = \{x: x \in A_t \textit{ for each member } t \textit{ of } T\}$. There are a number of elementary identities such as $\bigcup_{t \in T \cup S} A_t = (\bigcup_{t \in T} A_t) \cup (\bigcup_{t \in S} A_t)$, $C \setminus \bigcup_{t \in T} A_t = \bigcap_{t \in T} (C \setminus A_t)$ for all sets C (the de Morgan law), and $\bigcup_{t \in T} (B \cap A_t) = B \cap \bigcup_{t \in T} A_t$.

FUNCTIONS

We write $f: X \to Y$, which we read as "f is on X to Y", iff f is a map of X into Y; that is, f is a function with domain X whose values belong to Y. The value of the function f at a member x of X is denoted $f(x)$, or sometimes f_x.

If $f: X \to Y$ then "$x \mapsto f(x)$, for x in X", is another name for f. Thus $x \mapsto x^2$, for x in $\mathbb{R}$ (the set of real numbers) is the function that sends each real number into its square. The letter "x", in "$x \mapsto x^2$ for x in $\mathbb{R}$" is a dummy variable, so $x \mapsto x^2$ for x in $\mathbb{R}$ is the same as $t \mapsto t^2$ for t in $\mathbb{R}$. (Technically, "$\mapsto$" binds the variable that precedes it.)

If $f: X \to Y$ and $g: Y \to Z$ then $g \circ f: X \to Z$, the **composition** of g and f, is defined by $g \circ f(x) = g(f(x))$ for all x in X.

If $f: X \to Y$ and $A \subset X$ then $f|A$ is the **restriction of f to A** (that is, $\{(x, y): x \in A \text{ and } y = f(x)\}$) and $f[A]$ is the **image of A under f** (that is, $\{y: y = f(x) \text{ for some } x \text{ in } A\}$). If $B \subset Y$ then $f^{-1}[B] = \{x: f(x) \in B\}$ is the **pre-image** or **inverse image of B under f**. For each x, $f^{-1}[x]$ is $f^{-1}[\{x\}]$.

COUNTABILITY

A set A is **countably infinite** if there is a one to one correspondence between A and the set $\mathbb{N}$ of natural numbers (positive integers), and a set is **countable** iff it is countably infinite or finite.

Here is a list of the propositions on countability that we will use, with brief indications of proofs.

A subset of a countable set is countable.

If A is a subset of $\mathbb{N}$, define a function recursively by letting $f(n)$ be the first member of $A \setminus \{x: x = f(m) \text{ for some } m, m < n\}$. Then $f(n) \geqq n$ for each member n of the domain of f, and A is countably infinite if the domain of f is $\mathbb{N}$ and is finite otherwise.

The image of a countable set under a map is countable.

If f is a map of $\mathbb{N}$ onto A and $D = \{n: n \in \mathbb{N} \text{ and } f(m) \neq f(n) \text{ for } m < n\}$ then $f|D$ is a one to one correspondence between A and a subset of $\mathbb{N}$.

The union of a countable number of countable sets is countable.

It is straightforward to check that the union of a countable number of finite sets is countable, and $\mathbb{N} \times \mathbb{N}$ is the union, for k in $\mathbb{N}$, of the finite sets $\{(m, n): m + n = k + 1\}$.

If A is an uncountable set of real numbers then for some positive integer n the set $\{a: a \in A$ and $|a| > 1/n\}$ is uncountable.

Otherwise A is the union of countably many countable sets.

The family of all finite subsets of a countable set is countable.

For each n in $\mathbb{N}$, the family A_n of all subsets of $\{1, \ldots\ldots, n\}$ is finite, whence $\bigcup_n A_n$ is countable.

The family of all subsets of $\mathbb{N}$ is not countable.

If f is a function on $\mathbb{N}$ onto the family of all subsets of $\mathbb{N}$, then for some positive integer p, $f(p) = \{n: n \notin f(n)\}$. If $p \in f(p)$ then $p \in \{n: n \notin f(n)\}$, whence $p \notin f(p)$. If $p \notin f(p)$ then $p \notin \{n: n \notin f(n)\}$, whence $p \in f(p)$. In either case there is a contradiction.

ORDERINGS AND LATTICES

A relation $\geqq$ **partially orders** a set X, or **orders** X iff it is reflexive on X ($x \geqq x$ if $x \in X$) and transitive on X (if x, y and z are in X, $x \geqq y$ and $y \geqq z$ then $x \geqq z$). A **partially ordered set** is a set X with a relation $\geqq$ that partially orders it (formally, $(x, \geqq)$ is a partially ordered set). A member u of a partially ordered set X is an **upper bound** of a subset Y of X iff $u \geqq y$ for all y in Y; and if there is an upper bound s for Y such that $u \geqq s$ for every upper bound u of Y, then s is a **supremum** of Y, *sup* Y. A **lower bound** for Y and an **infimum** of Y, *inf* Y are defined in corresponding fashion.

An ordered set X is **order complete** or **Dedekind complete** iff each non-empty subset of X that has an upper bound has a supremum, and this is the case iff each non-empty subset that has a lower bound has an infimum.

A **lattice** is a partially ordered set X such that $\{x, y\}$ has a *unique* supremum and a *unique* infimum for all x and y in X. We denote $sup\{x, y\}$ by $x \vee y$ and $inf\{x, y\}$ by $x \wedge y$. A **vector lattice** is a vector space E over the set $\mathbb{R}$ of real numbers which is a lattice under a partial ordering with the properties: for x and y in E and r in $\mathbb{R}^+$ (the set of non-negative real numbers), if $x \geqq 0$ then $rx \geqq 0$, if $x \geqq 0$ and $y \geqq 0$ then $x + y \geqq 0$, and $x \geqq y$ iff $x - y \geqq 0$. Here are some properties of vector lattices:

For all x and y, $x \vee y = -((-x) \wedge (-y))$ and $x \wedge y = -((-x) \vee (-y))$, because multiplication by -1 is order inverting.

For all x, y and z, $(x \vee y) + z = (x + z) \vee (y + z)$ *and* $(x \wedge y) + z = (x + z) \wedge (y + z)$, because the ordering is translation invariant (i.e., $x \geqq y$ iff $x + z \geqq y + z$).

For all x and y, $x + y = x \vee y + x \wedge y$ (replace z by $-x - y$ in the preceding and rearrange).

If $x^+ = x \vee 0$ *and* $x^- = -(x \wedge 0) = (-x) \vee 0$ *then* $x = x \vee 0 + x \wedge 0 = x^+ - x^-$.

For each member x of a vector lattice E, the **absolute value of** x is defined to be $|x| = x^+ + x^-$. Vectors x and y are **disjoint** iff $|x| \wedge |y| = 0$.

For each vector x, x^+ *and* x^- *are disjoint,* because $x^+ \wedge x^- + x \wedge 0 = (x^+ + x \wedge 0) \wedge (x^- + x \wedge 0) = (x^+ - x^-) \wedge 0 = x \wedge 0$, whence $x^+ \wedge x^- = 0$.

The absolute value function $x \mapsto |x|$ completely characterizes the vector lattice ordering because $x \geqq 0$ iff $x = |x|$. On the other hand, if E is a vector space over $\mathbb{R}$, $A: E \to E$, $A \circ A = A$, A is **absolutely homogeneous** (i.e., $A(rx) = |r| A(x)$ for r in $\mathbb{R}$ and x in E), and A is additive on $A[E]$ (i.e., $A(A(x) + A(y)) = A(x) + A(y)$ for x and y in E), then E is a vector lattice and A is the absolute value, provided one defines $x \geqq y$ to mean $A(x - y) = x - y$.

(*Decomposition lemma*) *If* $x \geqq 0, y \geqq 0, z \geqq 0$ *and* $z \leqq x + y$, *then* $z = u + v$ *for some u and v with* $0 \leqq u \leqq x$ *and* $0 \leqq v \leqq y$. Indeed, we may set $u = z \wedge x$ and $v = z - z \wedge x$, and it is only necessary to show that $z - z \wedge x \leqq y$. But by hypothesis, $y \geqq z - x$ and $y \geqq 0$, so $y \geqq (z - x) \vee 0$, and a translation by $-z$ then shows that $y - z \geqq (-x) \vee (-z) = -(z \wedge x)$ as desired.

A real valued linear functional f on a vector lattice E is called **positive** iff $f(x) \geqq 0$ for $x \geqq 0$. If f is a positive linear functional, or if f is the difference of two positive linear functionals, then $\{f(u): 0 \leqq u \leqq x\}$ is a bounded subset of $\mathbb{R}$ for each $x \geqq 0$.

If f is a linear functional on E such that $f^+(x) = \sup\{f(u): 0 \leqq u \leqq x\} < \infty$ *for all* $x \geqq 0$, *then f is the difference of two positive linear functionals,* for the following reasons. The decomposition lemma implies that $\{f(z): 0 \leqq z \leqq x + y\} = \{f(u) + f(v): 0 \leqq u \leqq x \text{ and } 0 \leqq v \leqq y\}$, consequently f^+ is additive on $P = \{x: x \in E \text{ and } x \geqq 0\}$, and evidently f^+ is absolutely homogeneous. It follows that if x, y, u and v belong to P and $x - y = u - v$, then $f^+(x) - f^+(y) = f^+(u) - f^+(v)$, and f^+ can be extended to a linear functional on E—which we also denote by f^+. Moreover, $f^+ - f$ is non-negative on P and so $f = f^+ - (f^+ - f)$ is the desired representation.

The class E^* of differences of positive linear functionals on E is itself ordered by agreeing that $f \geqq g$ iff $f(x) \geqq g(x)$ for all x in E with $x \geqq 0$. Then E^*, with this ordering, is a vector lattice and $f^+ = f \vee 0$. It is to be emphasized that "f is positive" does *not* mean that $f(x) \geqq 0$ for *all* x in E, but only for members x of E with $x \geqq 0$.

Suppose a vector space F of real valued functions on a set X is ordered by agreeing that $f \geqq 0$ iff $f(x) \geqq 0$ for *all* x in X. If F, with this ordering, is a lattice, then it is a vector lattice and is called a **vector**

function lattice. This is equivalent to requiring that $(f \vee g)(x) = \max\{f(x), g(x)\}$ for all x in X.

CONVERGENCE IN $\mathbb{R}^*$

A relation $\geqq$ **directs** a set D iff $\geqq$ orders D and for each α and β in D there is γ in D such that $\gamma \geqq \alpha$ and $\gamma \geqq \beta$. Examples: the usual notion of greater than or equal to directs $\mathbb{R}$, the family of finite subsets of any set X is directed by $\supset$ and also by $\subset$, and the family of infinite subsets of $\mathbb{R}$ is directed by $\supset$ but not by $\subset$.

A **net** is a pair $(x, \geqq)$ such that x is a function and $\geqq$ directs the domain D of x. We sometimes neglect to mention the order and write the net x, or the net $\{x_\alpha\}_{\alpha \in D}$. A net with values in a metric space X (or a topological space) converges to a member c of X iff $\{x_\alpha\}_{\alpha \in D}$ is *eventually* in each neighborhood U of c; that is, if for each neighborhood U of c there is α in D such that $x_\beta \in U$ for all $\beta \geqq \alpha$. If $\{x_\alpha\}_{\alpha \in D}$ converges to c and to no other point, then we write $\lim_{\alpha \in D} x_\alpha = c$.

A **finite sequence** $\{x_k\}_{k=1}^n$ is a function on a set of the form $\{1, 2, \ldots, n\}$, for some n in $\mathbb{N}$. A **sequence** is a function on the set of positive integers, and the usual ordering of $\mathbb{N}$ makes each sequence a net. A sequence $\{x_n\}_{n \in \mathbb{N}}$ will also be denoted by $\{x_n\}_{n=1}^{\infty}$ or just by $\{x_n\}_n$. Thus for each q, $\{p + q^2\}_p$ is the sequence $p \mapsto p + q^2$ for p in $\mathbb{N}$.

It is convenient to extend the system of real numbers. The set $\mathbb{R}$, with two elements ∞ and $-\infty$ adjoined, is the **extended set** $\mathbb{R}^*$ of real numbers and members of $\mathbb{R}^*$ are **real*** numbers. We agree that ∞ is the largest member of $\mathbb{R}^*$, $-\infty$ is the smallest, and for each r in $\mathbb{R}$ we agree that $r + \infty = \infty + r = \infty$, $r + -\infty = -\infty + r = -\infty$, $r \cdot \infty = \infty$ if $r > 0$, $r \cdot \infty = -\infty$ if $r < 0$, $r \cdot (-\infty) = (-r) \cdot \infty$ for $r \neq 0$, $0 \cdot \infty = 0 \cdot (-\infty) = 0$, $\infty \cdot \infty = (-\infty) \cdot (-\infty) = \infty$ and $\infty \cdot (-\infty) = (-\infty) \cdot \infty = -\infty$.

Every non-empty subset of $\mathbb{R}$ which has an upper bound has a smallest upper bound, or supremum, in $\mathbb{R}$ and it follows easily that every subset of $\mathbb{R}^*$ has a supremum in $\mathbb{R}^*$ and also an infimum. In particular, $\sup \varnothing = -\infty$ and $\inf \varnothing = -\infty$.

A **neighborhood** in $\mathbb{R}^*$ of a member r of $\mathbb{R}$ is a subset of $\mathbb{R}^*$ containing an open interval about r. A subset V of $\mathbb{R}^*$ is a **neighborhood** of ∞ iff for some real number r, V contains $\{s\colon s \in \mathbb{R}^* \text{ and } s > r\}$. Neighborhoods of $-\infty$ are defined in a corresponding way. Consequently a net $\{x_\alpha\}_{\alpha \in D}$ in R^* converges to ∞ iff for each real number s there is β in D such that $x_\alpha > s$ for $\alpha \geqq \beta$.

If $\{x_\alpha\}_{\alpha \in A}$ and $\{y_\alpha\}_{\alpha \in A}$ are convergent nets in R^* then $\lim_{\alpha \in A}(x_\alpha + y_\alpha) = \lim_{\alpha \in A} x_\alpha + \lim_{\alpha \in A} y_\alpha$, provided the sum of the limits is defined and $\lim_{\alpha \in A} x_\alpha y_\alpha = (\lim_{\alpha \in A} x_\alpha) \cdot (\lim_{\alpha \in A} y_\alpha)$ provided the pair $(\lim_{\alpha \in A} x_\alpha, \lim_{\alpha \in A} y_\alpha)$ is not one among $(0, \pm\infty)$ or $(\pm\infty, 0)$. The proofs parallel those for nets in $\mathbb{R}$ with minor modifications.

If a net $\{x_\alpha\}_{\alpha \in A}$ in $\mathbb{R}^*$ is **increasing** (more precisely **non-decreasing**) in the sense that $x_\beta \geqq x_\alpha$ if $\beta \geqq \alpha$, then $\{x_\alpha\}_{\alpha \in A}$ converges to $sup_{\alpha \in A}\, x_\alpha$; for if $r < sup_{\alpha \in A}\, x_\alpha$, then r is not an upper bound for $\{x_\alpha\}_{\alpha \in A}$, consequently $r < x_\alpha$ for some α, and hence $r < x_\beta \leqq sup_{\alpha \in A}\, x_\alpha$ for $\beta \geqq \alpha$. Likewise, a decreasing net in R^* converges to $inf_{\alpha \in A}\, x_\alpha$ in R^*.

If $\{x_\alpha\}_{\alpha \in A}$ is a net in $\mathbb{R}^*$ then $\alpha \mapsto sup\{x_\beta: \beta \in A \text{ and } \beta \geqq \alpha\}$ is a decreasing net and consequently converges to a member of $\mathbb{R}^*$. This member is denoted $\mathbf{lim\,sup}_{\alpha \in A}\, x_\alpha$ or $\mathbf{lim\,sup}\{x_\alpha: \alpha \in A\}$. Similarly $\mathbf{lim\,inf}\{x_\alpha: \alpha \in A\}$ is $lim_{\alpha \in A}\, inf\{x_\beta: \beta \geqq \alpha\}$. It is easy to check that a net $\{x_\alpha\}_{\alpha \in A}$ converges iff $lim\,sup_{\alpha \in A}\, x_\alpha = lim\,inf_{\alpha \in A}\, x_\alpha$, and that in this case $lim_{\alpha \in A}\, x_\alpha = lim\,sup_{\alpha \in A}\, x_\alpha = lim\,inf_{\alpha \in A}\, x_\alpha$.

If $\{f_\alpha\}_{\alpha \in A}$ is a net of functions on a set X to $\mathbb{R}^*$ then $\mathbf{sup}_{\alpha \in A}\, f_\alpha$ is defined to be the function whose value at x is $sup_{\alpha \in A}\, f_\alpha(x)$, and similarly, $(\mathbf{inf}_{\alpha \in A}\, f_\alpha)(x) = inf_{\alpha \in A}\, f_\alpha(x)$, $(\mathbf{lim\,sup}_{\alpha \in A}\, f_\alpha)(x) = lim\,sup_{\alpha \in A}\, f_\alpha(x)$ and $(\mathbf{lim\,inf}_{\alpha \in A}\, f_\alpha)(x) = lim\,inf_{\alpha \in A}\, f_\alpha(x)$. The net $\{f_\alpha\}_{\alpha \in A}$ **converges pointwise** to f iff $f = lim\,sup_{\alpha \in A}\, f_\alpha = lim\,inf_{\alpha \in A}\, f_\alpha$ or, equivalently, $f(x) = lim_{\alpha \in A}\, f_\alpha(x)$ for all x.

UNORDERED SUMMABILITY

Suppose $x = \{x_t\}_{t \in T}$ is an indexed family of real* numbers. We agree that $\{x_t\}_{t \in T}$ is **summable* over a finite subset** A of T iff x does not assume both of the values ∞ and $-\infty$ at members of A, and in this case the sum of x_t for t in A is denoted by $\sum_{t \in A} x_t$ or $\sum_A x$. If $\{x_t\}_{t \in T}$ is summable* over each finite subset, and if $\mathscr{F}$ is the class of all finite subsets of T, then $\mathscr{F}$ is directed by $\supset$, $\{\sum_A x\}_{A \in \mathscr{F}}$ is a net, and we say that x is **summable* over** T, or just **summable*** provided that the net $\{\sum_A x\}_{A \in \mathscr{F}}$ converges. In this case the **unordered sum**, $\sum_T x$, is $lim\{\sum_A x: A \in \mathscr{F}\}$, and $\{x_t\}_{t \in T}$ is summable* to $\sum_T x$.

If $x = \{x_t\}_{t \in T}$ is a family of real numbers, then x is automatically summable* over each finite subset of T and we say that x is **summable over** T, or just **summable**, provided it is summable* and $\sum_T x \in \mathbb{R}$.

If $\{x_n\}_{n \in \mathbb{N}}$ is a sequence of real numbers then the (ordered) sum, $lim_n \sum_{k=1}^n x_k$, may exist although the sequence is not summable (e.g., $x_n = (-1)^n/n$ for each n in $\mathbb{N}$). However, if $\{x_n\}_n$ is summable* then the limit of $\{\sum_{k=1}^n x_k\}_n$ exists and $lim_n \sum_{k=1}^n x_k = \sum_{n \in \mathbb{N}} x_n$.

Here are the principal facts about unordered summation, with a few indications of proof. Throughout, $x = \{x_t\}_{t \in T}$ and $y = \{y_t\}_{t \in T}$ will be indexed families of real* numbers, $(x^+)_t = (x_t)^+$ and $(x^-)_t = (x_t)^-$ for each t, and r will be a real number.

The family $x = \{x_t\}_{t \in T}$ *is summable iff for* $e > 0$ *there is a finite subset* A *of* T *such that* $\sum_B |x| < e$ *for each finite subset* B *of* $T \setminus A$.

If $x = \{x_t\}_{t \in T}$ *is summable then* $x_t = 0$ *except for countably many points* t.

If $x_t \geqq 0$ *for each* t *then* $\{x_t\}_{t \in T}$ *is summable**.

(The net $\{\sum_A x : A \in \mathscr{F}\}$ is increasing.)

The family x is summable iff one of* $\sum_T x^+$ *and* $\sum_T x^-$ *is finite; it is summable iff both are finite; and in either of these two cases,* $\sum_T x = \sum_T x^+ - \sum_T x^-$. (The result reduces to the usual "limit of the difference" proposition.)

If x is summable and* $r \in \mathbb{R}$ *then rx is summable* and* $\sum_T rx = r\sum_T x$.

The next proposition states that "$\sum_X$ is additive except for $\infty - \infty$ troubles". It's another "limit of a sum" result.

If x and y are summable,* $\{x_t, y_t\} \neq \{\infty, -\infty\}$ *for all t, and* $\{\sum_T x, \sum_T y\} \neq \{\infty, -\infty\}$, *then* $x + y$ *is summable* and* $\sum_T (x + y) = \sum_T x + \sum_T y$.

If x is summable over T and* $A \subset T$ *then x is summable* over A.*

If x is summable over T and* $\mathscr{B}$ *is a disjoint finite family of subsets of T then* $\sum_{B \in \mathscr{B}} (\sum_B x) = \sum \{x_t : t \in \bigcup_{B \in \mathscr{B}} B\}$.

If $\mathscr{A}$ *is a decomposition of T (i.e., a disjoint family of subsets such that* $T = \bigcup_{A \in \mathscr{A}} A$) *and x is summable* over T then* $A \mapsto \sum_A x$ *is summable* over* $\mathscr{A}$ *and* $\sum_T x = \sum_{A \in \mathscr{A}} \sum_A x$.

If x is summable over* $Y \times Z$, *then* $\sum_{Y \times Z} x = \sum_{y \in Y} \sum_{z \in Z} x(y, z) = \sum_{z \in Z} \sum_{y \in Y} x(y, z)$.

It is worth noticing that the condition, "*x is summable**", is necessary for the last equality. Here is an example. Define x on $\mathbb{N} \times \mathbb{N}$ by letting $x(m, n)$ be 1 if $m = n$, -1 if $n = m + 1$, and 0 otherwise. Then $\sum_{m \in \mathbb{N}} x(m, n) = 0$ if $n > 0$ and 1 if $n = 0$, so $\sum_{n \in \mathbb{N}} (\sum_{m \in \mathbb{N}} x(m, n)) = 1$, whereas $\sum_{m \in \mathbb{N}} (\sum_{n \in \mathbb{N}} x(m, n)) = \sum_{m \in \mathbb{N}} (0) = 0$.

A family $\{f_t\}_{t \in T}$ of real* valued functions on a set X is **pointwise summable*** (**summable**, respectively) iff $\{f_t(x)\}_{t \in T}$ is summable* (summable, respectively) for each x in X, and in this case the **pointwise sum**, $(\sum_{t \in T} f_t)(x)$ is defined to be $\sum_{t \in T} f_t(x)$ for each x in X.

HAUSDORFF MAXIMAL PRINCIPLE

If $\geqq$ partially orders X then a subset C of X is a **chain** iff for all x and y in C with $x \neq y$, either $x \geqq y$ or $y \geqq x$ but not both. We assume (and occasionally use) the following form of the maximal principle.

ZORN'S LEMMA *If C is a chain in a partially ordered space* $(X, \geqq)$ *then C is contained in a maximal chain D—that is a chain that is a proper subset of no other chain.*

Consequently, if every chain in X has a supremum in X then there is a maximal member m of X—that is, *if* $n \geqq m$ *then* $n = m$.

Here is a simple example of the application of the maximal principle. Suppose that G is a subset of the real plane $\mathbb{R}^2$ and that $\mathscr{D}$ is the family of disks $D_r(a, b) = \{(x, y) : (x - a)^2 + (y - b)^2 \leqq r^2\}$ with (a, b) in $\mathbb{R}^2$, $r > 0$ and $D_r(a, b) \subset G$. Then there is a maximal disjoint subfamily $\mathscr{M}$ of $\mathscr{D}$, and $G \setminus \bigcup_{D \in \mathscr{M}} D$ contains no non-empty open set.

Chapter 1

PRE-MEASURES

We consider briefly the class of length functions. These will turn out to be precisely the functions on the family of closed intervals that can be extended to become measures; these are examples of pre-measures. Their theory furnishes a concrete illustration of the general construction of measures.

A **closed interval** is a set of the form $[a:b] = \{x: x \in \mathbb{R} \text{ and } a \leqq x \leqq b\}$, an **open interval** is a set of the form $(a:b) = \{x: a < x < b\}$, and $(a:b]$ and $[a:b)$ are half open intervals. The family of closed intervals is denoted $\mathscr{I}$; we agree that $\varnothing \in \mathscr{I}$. We are concerned with real valued functions λ on $\mathscr{I}$, and we abbreviate $\lambda([a:b])$ by $\lambda[a:b]$. The closed interval $[b:b]$ is just the singleton $\{b\}$, and $\lambda[b:b] = \lambda(\{b\})$ is abbreviated $\lambda\{b\}$.

A non-negative real valued function λ on $\mathscr{I}$ such that $\lambda(\varnothing) = 0$ is a **length**, or a **length function for R**, iff λ has three properties:

Boundary inequality If $a < b$ then $\lambda[a:b] \geqq \lambda\{a\} + \lambda\{b\}$.
Regularity If $a \in \mathbb{R}$ then $\lambda\{a\} = \inf\{\lambda[a-e:a+e]: e > 0$.
Additive property If $a \leqq b \leqq c$ then $\lambda[a:b] + \lambda[b:c] = \lambda[a:c] + \lambda[b:b]$.

The **length**, or the **usual length function** ℓ, is defined by $\ell[a:b] = b - a$ for $a \leqq b$. The length ℓ is evidently a length function; it has a number of special properties—for example, $\lambda\{x\} = 0$ for all x.

There are length functions that vanish except at a singleton. The **unit mass at a member** c of $\mathbb{R}$, ε_c, is defined by letting $\varepsilon_c[a:b]$ be one if $c \in [a:b]$ and zero otherwise. Thus $\varepsilon_c\{x\} = 0$ if $x \neq c$ and $\varepsilon_c\{c\} = 1$. Each such unit mass is a length function, and each non-negative, finite linear combination of unit masses is a length function.